Measuring the Effect of the Timing of First Birth*

Jane Leber Herr†

August 2012

Abstract

I study the effect of first-birth timing on women's wages, defining timing in terms of labor force entry, rather than age. Considering the mechanisms by which timing may affect wages, each is a function of experience rather than age. This transformation also highlights the distinction between a first birth after labor market entry versus before. I show that estimates based on age understate the return to delay for women who remain childless at labor market entry, and have obscured the *negative* return to delay – to a first birth *after* labor market entry – for all but college graduates. **[J13, J31]**

*I would like to thank David Card, Ronald Lee, Guido Imbens, Jesse Shapiro, Derek Neal, Elizabeth Weber Handwerker, Melanie Guldi, Emily Oster, Colleen Manchester, Sarah Hayford, Lucie Schmidt, Don Cox, Jenna Johnson-Hanks, Ashley Langer, Kevin Stange, Donna Ginther, Martha Bailey, Rachel Sullivan Robinson and participants of the U.C. Berkeley labor seminar series, the University of Chicago's Booth applied microeconomic series, and the Federal Reserve Bank of Chicago's seminar series for helpful comments. I would also like to thank Jay Zagorsky and Steve McClaskie from the Center of Human Resource Research at the Ohio State University for providing me the sampling weights appropriate for my sample. Lastly, I would like to thank the Center for Human Potential and Public Policy at the University of Chicago's Harris School, and its director, Ariel Kalil, for support and guidance. Financial support for this work was provided by the National Institute for Child Health and Human Development (Interdisciplinary Training Grant No. T32-HD007275).

†The National Bureau of Economic Research, 1050 Massachusetts Avenue, Cambridge, MA 02138, herrj@nber.org.

1 Introduction

Over the last several decades, a rich vein of economic research has established that women with children receive lower wages than those without – the so-called "motherhood wage gap" (see, for instance, Waldfogel 1998). A related area of research has focused on whether the timing of entry into motherhood influences the magnitude of this cost. Although research has shown that women who delay entry into motherhood have systematically higher long-run wages (Bloom 1986), identifying the causal effect of first-birth timing is complicated by women's capacity to (imperfectly) control their fertility.

Although the existing literature has consistently found positive bias captured in the correlation between first-birth timing and wages, it has also concluded that there remains a benefit to fertility delay. Chandler *et al.* (1994) and Miller (2011) find that a one-year delay of first birth is associated with 2 to 3 percent higher wages, measured among women in their 30s or 40s. Other research has found that a first birth beyond age 28 or 30 closes the motherhood wage gap (Taniguchi 1999; Amuedo-Dorantes and Kimmel 2005).

A key limitation of this research, however, is its focus on *age* at first birth. As I show in this paper, to gauge the effect of fertility timing on wages, the appropriate measure is instead a woman's "career timing" of first birth, the point in her labor market career in which children are first present. Furthermore, I show that the existing literature's focus on age has produced confused inference that obscures the magnitude, and for some the *sign*, of the link between timing and wages.

To set the stage for assessing the effect of first-birth timing, I begin by considering the mechanisms by which timing may affect a woman's wage path. This thought experiment reveals a first key insight: each potential mechanism turns on a woman's *experience* level at first birth, not her age. I therefore begin my analysis by translating age at first birth into a woman's "relative timing" – the timing of her first birth relative to when she enters the labor force.

This transformation then reveals a second key insight: the relative timing can only reflect a woman's experience level at first birth for those who have their first child *after* they enter the labor force. For women who instead have a first birth before working, the mechanism by which fertility delay may affect subsequent wages is more ambiguous. Thus one main contribution of this paper is the recognition that the link between first-birth timing and wages may be different for these two populations of mothers – those with a first birth before entering the labor force, and those with a first birth afterwards.

Formalizing these thoughts into a model of the effect of first-birth timing on a woman's wage path, I then show the confused inference that may arise by instead measuring timing in terms of age, and in doing so, combining these two groups. First, I show that the coefficient estimate on age is likely to *under*-state the return to delay for those women who remain childless at labor market entry. And second, I show that this coefficient will obscure the relationship between timing and wages for women who instead enter the labor force after motherhood, and may in fact indicate the wrong sign.

I next turn to the question of the potential bias captured in the raw correlation between first-birth timing and wages. Using the wage model described above as a starting point, I formalize the economic conditions under which bias may be captured in the estimated return to delay. Focusing first on the population who enter motherhood after labor market entry – the majority of women – I show that each source of bias discussed in the literature makes clear predictions for how wages diverge between early and late mothers.

I then return to the question of how women jointly choose (*i*) their year of first birth and (*ii*) their year of labor market entry, the *combination* of which sorts them into these two populations of mothers. Using women from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79), although I find evidence suggesting that the former choice is primarily driven by variation in taste for early motherhood, I find that the latter is clearly related to a woman's potential wage. Yet when I plot observable characteristics by "relative timing," there are no discontinuities across the threshold between these two populations of mothers. This suggests that my use of the timing of labor market entry to translate first birth timing into a measure of experience rather than age likely has not created a new source of bias in my estimates.

As a first gauge of the relationship between first-birth timing and wages, I then plot the wage path of women from the NLSY79, grouped by their relative timing. Whereas I find a clear relationship between timing and wages for women who remain childless at labor market entry, there is no such link for women who instead enter the labor market after motherhood. Furthermore, comparing the former to the predicted wage paths associated with key potential sources of bias, I find that the raw wage pattern is strikingly similar to the wage path consistent with an unbiased reflection of the causal effect of fertility timing.

I follow this with a simple OLS framework that considers how the measured relationship between first-birth timing and wages varies as one controls for an increasing number of factors that are correlated with either (*i*) wage growth, (*ii*) taste for early motherhood, or (*iii*) the individual return to fertility delay. In this exercise I consider the effect of these controls on the estimated slope between timing and wages for both populations of mothers. For women who remain childless at labor market entry, I again find surprisingly little evidence of bias captured in the positive correlation between timing and wages. For women with a first birth before working, I instead find that the observed positive correlation is the result of heterogeneity.

Overall, my results show that there is a clear benefit to continued fertility delay for women who enter the labor force before motherhood. For those with at least a high school diploma when they start working, a year of fertility delay leads to 1.5 to 3.6 percent higher wages 20 years later. By contrast, I find that there is no *linear* relationship between fertility timing and wages for women who have their first child before they start working. I therefore conclude that the relevant measure of the timing of a woman's first birth is her "career timing," the point in her career when children are first present.

Yet counter to expectations, for all but college graduates, I also find evidence suggesting that many women who had their first child *after* labor market entry, would have had higher long-run wages if they instead had their first birth earlier, *before* entering the labor force. Thus I find an initial *negative* return to fertility delay – at least across the threshold at labor market entry – followed by a positive return to subsequent delay further into one's career.

The remainder of this paper is structured as follows. Section 2 discusses the rationale for measuring fertility timing in terms of the point of labor market entry, and lays out a corresponding model of the effect of first-birth timing on a woman's long-run wage level. In Section 3 I then discuss the key sources of potential bias that may be captured in the correlation between first-birth timing and wages. In Section 4 I introduce my sample and variable definitions, followed by a discussion of my identification strategy in Section 5. Section 6 reports my results, and Section 7 concludes.

2 Modeling the Effect of First-Birth Timing on Wages

In this section I begin by discussing the rationale for measuring the timing of first birth in terms of the year that a woman enters the labor market. I then propose a model for the effect of first-birth timing on a woman's long-run wage level, ending with a discussion of the specification error that may arise if timing is instead measured in terms of age.

2.1 Measurement of the Timing of First Birth

To gauge the appropriate measure of fertility timing, consider the mechanisms by which timing may affect a woman's wages. In the wage profile of male workers, wages rise quickly over a worker's first decade in the labor market, grow at an increasingly more moderate pace in the subsequent decade, and slowly flatten out from that point forward (Murphy and Welch 1990). Taking this as the potential wage path for a woman entering the labor force at age 20, we see that the steepest period of wage growth falls during her key childbearing years, and that wages do not fully plateau until after her childbearing years have passed.

Since wage growth is steeper earlier in one's career, one can imagine that a woman faces a payoff to delaying her first birth – to putting off the discontinuous shock to time demands that arise at motherhood. More formally, if the timing of a woman's first birth affects her wages, it most likely does so through its influence on those economic behaviors that generate the shape of the standard wage path.[1]

For instance, if a key factor is an especially high rate of return on early-career work experience (Light and Ureta 1995), the return to fertility delay may arise from the payoff to one more year of full-intensity work experience during this initial stage.[2] Fertility timing may also influence wages through its effect on a woman's labor supply, if the greater human capital accumulated by the time of a delayed first birth lowers her propensity to leave the labor force at motherhood, or the length of time she takes off. Alternatively, following job match theory (Topel and Ward 1992), if the arrival of a baby discontinuously increases the transaction costs of job search, fertility delay may affect a woman's wage path by allowing her to reach a better-quality job match before the transition into motherhood.

Although the purpose of this paper is not to establish the mechanism by which fertility timing affects wages, this thought experiment highlights the fact that each of these possibilities centers on the timing of a woman's first birth in terms of experience, not age. I therefore begin my analysis by calculating a woman's "relative timing," K_1 – the difference between the year of her first birth and the year she enters the labor force, t_1.[3] Thus

[1]One might also ask whether the *reason* for fertility delay influences the return, if women's subsequent labor market behavior varies systematically with that reason. One can imagine isolating the factor that was the binding constraint the year before we see a woman get pregnant with her first child. For instance, "I would have gotten pregnant one year earlier, but (a) I wasn't married yet; (b) my husband's earnings weren't high enough yet; (c) it took me a long time to get pregnant; (d) I was up for a big promotion; (e) I didn't want a baby yet." (Or, in the reverse, (f) "I would have waited one more year, but I got pregnant on accident.") One can therefore think of the average return to a year of delay reflecting the weighted average of these individual returns. See footnote 56 for more on this question.

[2]For the vast majority of mothers, the arrival of a first child marks a clear drop in her labor force presence, whether or not she stops working. For instance, in the NLSY79, among women who are working full-time before motherhood, at least two thirds cut their hours by at least 10 percent in the following year, and only 20 percent do not reduce their labor supply (even when calculated only among women who remain working).

[3]Note that in practice, this raises an issue of potential measurement error, since a woman's year of labor market entry may not be clear cut. See Section 4 and Appendix Section C for both a detailed description of how I define t_1 relative to when a woman "completes" school, as well as a discussion of the possible endogeneity of the latter. (Also see footnote 20 for the results using a stricter definition of t_1.)

$K_1 > 0$ for the population who have their first child after they begin working, and $K_1 \leq 0$ for the population who enter the labor force with a child. And note that K_1 only reflects the "experience timing" of first birth for the former.[4]

2.2 Modeling the Effect of Timing

Given this argument, in the following section I begin by writing down a separate model of the potential effect of timing on wages for each population of mothers. First, consider the women who enter the labor market before having children ($K_1 > 0$). Suppose that their wages grow at a rate g_1 up till the year of their first birth, K_1, at which point wage growth stalls to a slower rate, g_2. This drop may arise, for instance, from a change in labor supply, or in effort per hour worked (Becker, 1985). Yet as noted in Section 2.1, since the change in labor supply at first birth may be endogenous to its timing – acting as a *mechanism* by which timing affects wages – I do not incorporate such labor supply behaviors directly into the model, in order to capture the full effect of fertility timing on wages.[5]

Thus, given a woman's starting hourly wage, w_0, the log of her wage level at some point $\tau > K_1$ will be a linear function of her first-birth timing,

$$l(w_\tau) = l(w_0) + \theta K_1 + g_2 \tau, \tag{1}$$

where the coefficient on K_1 is the change in her wage growth at that point, $\theta \equiv g_1 - g_2$.

Now consider the population who enter the labor market with children ($K_1 \leq 0$). For these pre-t_1 mothers ($pt_1 = 1$), the mechanism for the link between first-birth timing and the

[4]Although K_1 only reflects a woman's *potential* experience at first birth, in practice it reflects actual experience for the vast majority of these women. For instance, in the NLSY79, among women with $K_1 > 0$, 95 percent or more are in the labor force each year between t_1 and the year before first birth.

[5]In contrast, the models in Happel *et al.* (1984), Moffitt (1984), Cigno and Ermisch (1989) and Walker (1995) assume some exogenously-defined level of maternal time associated with child-raising. One might instead assume a temporary, rather than permanent, drop in the growth rate at motherhood, as in the models in Wilde *et al.*, 2010, and Mullin and Wang, 2002 (although the latter collapses the three-period model into a two-period version that assumes a permanent drop). I rely on the specification written here in light of Figure 5a, which shows no evidence of a return to the pre-birth growth rate, at least within the 10- to 15-year post-birth window observed.

subsequent wage path is more ambiguous. One possibility is that K_1 has a direct effect on the wage growth rate of these mothers, g_{pt_1}, if there is a link between the age of a woman's oldest child at t_1 (by definition equal to $-K_1$), and her subsequent labor market behavior: $g_{pt_1} = g_0 + g_0^{'}(-K_1)$. Combining this with Equation (1) provides a model of the effect of first-birth timing on wages for both populations of mothers:[6]

$$l(w_\tau) = l(w_0) + (1 - pt_1)\big(\theta K_1 + g_2\tau\big) + pt_1\Big((g_0 - g_0^{'}K_1)\tau\Big). \tag{2}$$

To show this visually, Figure 1 depicts the relationship between timing and the long-run wage level, focusing on different possible scenarios for pre-t_1 mothers ($K_1 \leq 0$). For instance, as in Figure 1a, there may be a negative relationship between timing and wages ($g_0^{'} > 0$), if, for instance, mothers who enter the labor force with older children work longer hours, leading to faster wage growth. Alternatively, there may be a positive slope if mothers with more children at t_1 – which is monotonically decreasing in K_1 – work shorter hours.

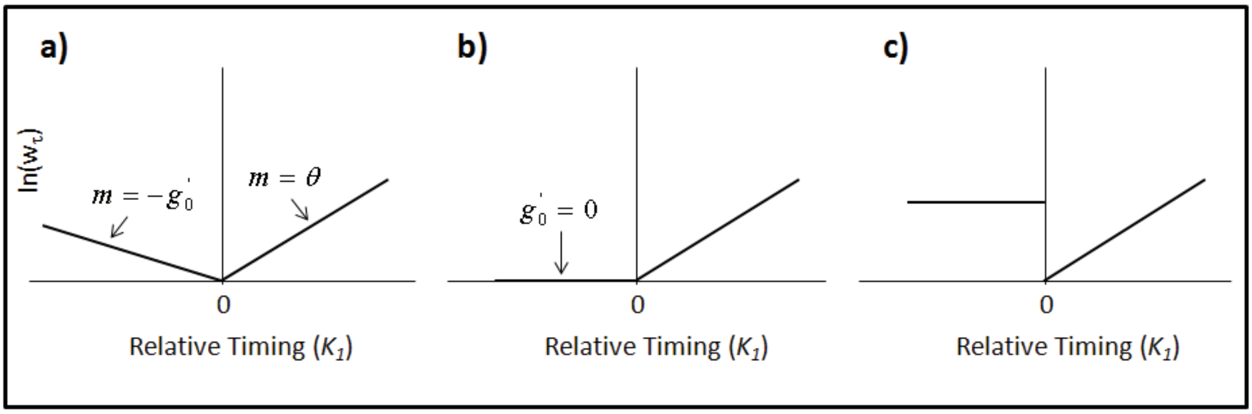

Figure 1: Possible Links Between First-Birth Timing and Long-Run Wages

Another possibility, depicted in Figures 1b and 1c, is that there is no direct link between timing and wages for pre-t_1 mothers, $g_0^{'} = 0$, thus $g_{pt_1} = g_0$. Furthermore, if, as drawn in Figure 1b, there is no trend break at $K_1 = 0$, this implies that $g_0 = g_2$, the post-motherhood growth rate for post-t_1 mothers ($K_1 > 0$). Thus under these circumstances it is strictly better to have one's first birth after labor market entry, rather than before. Yet if there is

[6]Note that as written in Equation 2, the starting wage (w_0) is unrelated to first birth timing. Yet for pre-t_1 mothers, one might anticipate that w_0 is a function of K_1, or at least of the presence of children. In Section 6 I therefore report my results both with and without controlling for starting wages.

a positive trend break, as in Figure 1c, then $g_0 > g_2$, and it is instead especially costly to interrupt one's career for motherhood shortly after entering the labor force. Thus in this scenario there is a negative return to delay, at least across the threshold at t_1.

Implications of Measuring Timing in Terms of Age

Now consider the implications of instead estimating the effect of timing on wages using age at first birth, a_{b_1}. Note that K_1 is simply a linear transformation of a_{b_1}: $K_1 = a_{b_1} - a_{t_1}$, where a_{t_1} is age at t_1. Thus rewriting Equation (2) to focus on K_1,

$$l(w_\tau) = l(w_0) + \Big((1 - pt_1)\theta - pt_1 g_0^{'}\tau\Big)K_1 + \Big((1 - pt_1)g_2 + pt_1 g_0\Big)\tau,$$

we see that a regression of $l(w_\tau)$ on a_{b_1} will capture an estimate of $\Big((1 - pt_1)\theta - pt_1 g_0^{'}\tau\Big)$.

If the sample used in such a regression includes only women with a first birth *after* labor market entry ($pt_1 = 0$), then the coefficient on a_{b_1} will recover the return to delay for this population, θ. In a mixed population of pre- and post-t_1 mothers, however, the extent to which this coefficient reflects an unbiased estimate of θ is unclear. Assuming $\theta > 0$, the coefficient on a_{b_1} will only recover the return to delay for both populations if the slope is the same on both sides of 0. And note that if this holds, it will not matter that the existing literature has grouped these two populations of mothers.[7]

But if instead $g_0^{'} \geq 0$, as the proportion of pre-t_1 mothers rises, the coefficient on a_{b_1} will increasingly underestimate θ, the return for women with $K_1 > 0$. Furthermore, since pre-t_1 mothers represent a minority of women, the coefficient on a_{b_1} will likely do an even poorer job of reflecting their return to delay, and may even reflect the wrong sign. Assuming θ is positive, the latter will clearly hold if $g_0^{'} > 0$. Yet even if $g_0^{'} = 0$, this will still hold if $g_0 > g_2$, as in Figure 1c, where there is a negative return to delay across the threshold at t_1. In sum, the limitations of the existing literature's focus on age at first birth stems primarily from its grouping of these two populations of mothers.

[7]This will only occur if $g_0^{'} < 0$, and if measured in the specific year in which $\theta = -g_0^{'}\tau$.

3 Potential Sources of Bias in Estimating the Effect of Timing

In the following section I now consider how various sources of bias may influence the observed relationship between first-birth timing and wages, given women's capacity to control their fertility, as well as their timing of labor market entry. Building on the wage model introduced above, in Section 3.1 I begin by modeling how women choose their optimal timing of first birth, K_1^*, focusing on women for whom the optimal timing is after labor market entry, t_1.

In Section 3.2 I then use this model to consider how several key potential sources of bias may affect the relationship between first-birth timing and wages, specifically highlighting their influence on the path of wages over time, as well as the correlation between timing and the long-run wage level. By limiting this discussion to women with $K_1 > 0$, I focus here on bias captured in the estimates of θ, the relevant parameter of the return to delay for the majority of women. Lastly, in Section 3.3, I return to the question of women's control over their timing of labor market entry, t_1 – a key building block of my measure of "relative timing," K_1 – and how this potential endogeneity may generate additional bias in my estimates of the effect of first-birth timing.

3.1 Modeling Optimal Timing, K_1^*

To begin, suppose women choose K_1^* to maximize the following lifetime utility function, $U = log(y) - c(K_1)$, where utility is increasing in own earnings, y, but decreasing with delayed first birth. Furthermore, assume earnings are a function of K_1 via the wage path in Equation (1), and that the cost of delay, $c(K_1)$, varies across women with their taste for early motherhood, ψ. (Note that as written, I ignore the question of partnering or marriage.[8])

To solve for a woman's marginal benefit of fertility delay, we must first solve for the production function of earnings, $y = f(K_1)$. Making the simplifying assumption that labor

[8]By comparison, Happel *et al.* (1984), Moffitt (1984), and Cigno and Ermisch (1989) treat the timing of first marriage as the start of the planning horizon (taking its timing, and a woman's human capital at that point, as exogenous). As here, in most of these models, lifetime utility is increasing in an earlier first birth.

supply (h_t) is constant throughout the lifecycle, the net present value of lifetime earnings is

$$f(K_1) = \int_T w_t h_t e^{-rt} dt = \int_0^{K_1} w_0 e^{g_1 t} e^{-rt} dt + \int_{K_1}^{T} w_0 e^{g_1 K_1} e^{g_2 (t-K_1)} e^{-rt} dt,$$

where r reflects the discount rate and T the length of a woman's career. Taking the linear approximation of $f'(K_1)/f(K_1)$ provides the marginal benefit of fertility delay, $MB(K_{1i}) = \theta_i - m_{1i} K_{1i}$, where $m_{1i} \geq 0$.[9] Thus for each woman, the marginal benefit of delaying her first birth by one more year has an intercept term equal to the change in her wage growth rate at first birth, $\theta_i = g_{1i} - g_{2i}$, and the return is decreasing in K_1.

I can similarly approximate the marginal cost of delay, $c'(K_1)$, as a linear function of a woman's taste for early motherhood, ψ: $MC(K_1) = \psi_i + m_{2i} K_{1i}$, where I assume $m_{2i} \geq 0$.[10] A woman will therefore choose K^*_{1i}, her optimal timing of first birth, at the point where the marginal benefit of one more year of delay is just equal to its marginal cost, solving to:

$$K^*_{1i} = \frac{\theta_i - \psi_i}{m_{1i} + m_{2i}}. \tag{3}$$

3.2 Potential Influence of Bias in Estimating θ

Building on this optimization model, I can now consider how various sources of bias may manifest themselves in the observed relationship between wages and first-birth timing. In particular, I consider four scenarios:

1. no bias captured in the correlation between timing and wages,
2. the correlation is influenced by endogenous timing of first birth,
3. the correlation arises only from heterogeneity in the population, rather than from a causal effect of timing, and
4. the link between timing and wages is driven by reverse causality.

These last three reflect the key sources of bias discussed in the existing literature.

[9]See Appendix Section A for the solution to $f(K_1)$ and my calculation of the marginal benefit of delay. In the appendix I also discuss the corresponding calculation if h_t varies over time.

[10]Since the marginal cost of delay will be a function of declining fertility, one could rewrite this in terms of age at labor market entry.

For each scenario, I build a simplified figure of the expected path of wages over time, comparing two women, one with an early, and the other with a late first birth. (A key consideration throughout is what generates the variation in optimal timing, K_1^*.) In Section 6 I will return to compare these figures to the raw path of wages observed for the women in the NLSY79, as a first gauge of the sources of bias that may be captured in the observed correlation between first-birth timing and long-run wages.

Scenario #1: No Bias

Consider a first possible scenario, in which the rate of wage growth before and after first birth, g_1 and g_2, are equal for all women ($g_{1i} = g_1$ and $g_{2i} = g_2 \ \forall \ i$). The return to delay is therefore invariant across the population, $\theta_i = \theta$ for all i. In this scenario, the marginal benefit of delay is thus constant, and from Equation (3) we see that variation in K_1^* will arise only through variation in taste for early motherhood, ψ.

The first panel of Figure 2 plots the path of wages under this scenario, comparing two women, one who chooses an early first birth, K_1^E, and the other who chooses a late first birth, K_1^L.[11] Assuming an equivalent starting wage, we see that their wages move together, growing at the rate g_1, up till the point when the 'early' mother has her first birth, $t = K_1^E$. Her wage path then begins to grow at the slower rate g_2, while the 'late' mother's wage continues to grow at g_1 until she has her first birth at $t = K_1^L$.

[11] Each panel of Figure 2 plots the path of wages over time, from labor market entry to some year τ beyond childbearing, for two women, one with an early first birth (K_1^E) and the other with a late first birth (K_1^L). In each, I assume that both women enter the labor market at the same starting wage, $l(w_0)$. In the first scenario, each experiences the same rate of wage growth before first birth and after, $g_1^E = g_1^L$ and $g_2^E = g_2^L$. In the second, the later mother has higher ability ($a^L > a^E$), and has thus chosen to delay first birth in response to a higher return to delay ($\theta = \theta(a), \theta' > 0$). She therefore experiences faster wage growth before first birth, $g_1^L > g_1^E$ (because $g_1 = g_1(a), g_1' > 0$), but by assumption g_2 remains equal. In the third scenario, there is no change in wage growth at first birth ($g_{1i} = g_{2i} = g_i \ \forall \ i$), but the later mother again has higher ability – in this case because $Cov(\psi, a) < 0$ – and $g^L > g^E$ because $g = g(a)$ with $g' > 0$. In the fourth scenario, women experience the same rate of early-career wage growth, $g_{1i}' = g_1' \ \forall \ i$, up till some critical point in their career, t_c (the timing of which varies randomly across the population), at which point each woman's wage growth falls to a lower rate, g_2' (where $g_{2i}' = g_2' \ \forall \ i$). This scenario also assumes that women systematically delay conceiving their first child until after they reach t_c, and that women take the same amount of time to conceive and give birth after that point.

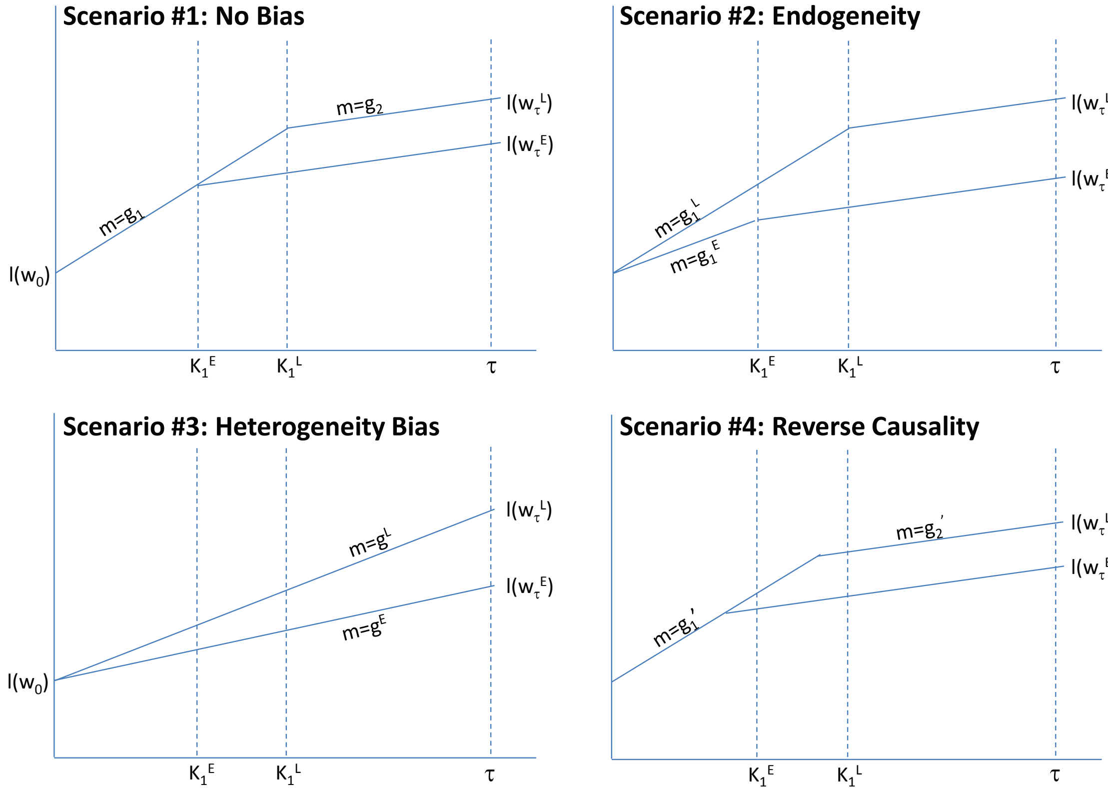

Figure 2: Possible Wage Paths by Timing of First Birth

If we then trace their wage path through some point τ beyond K_1^L, we see that the difference in their final wage level is equal to θ, the return to one more year of fertility delay,

$$l(w_\tau^L) - l(w_\tau^E) = \Big(l(w_0) + \theta K_1^L + \tau g_2\Big) - \Big(l(w_0) + \theta K_1^E + \tau g_2\Big) = \theta(K_1^L - K_1^E) = \theta,$$

if I make the simplifying assumption that $K_1^L - K_1^E = 1$. Under this scenario, running a least squares regression of $l(w_{\tau i})$ on observed timing of first birth, K_{1i}, will therefore provide an unbiased estimate of the average return to delay, $\bar{\theta} = \theta$.

Scenario #2: Endogenous Timing of First Birth

Now suppose that the observed correlation between timing and wages is influenced by endogenous timing of first birth in response to variation in the return to delay. For instance, suppose θ is an increasing function of ability, a, as in Mullin and Wang (2002). To translate this into the terms of Equation (1), suppose $g_{1i} = g_1(a_i)$, with $g'_1 > 0$ and g_2 constant.[12] In this case, via Equation (3) we see that higher-ability women choose later first births in response to their greater marginal benefit of delay.

The second panel of Figure 2 plots the wage path under this scenario of endogenous timing, where the two women now vary in ability, $a^E < a^L$. Maintaining the assumption that they start at the same wage level, we see that because $g_1^E < g_1^L$, their wages diverge *before* either have children. If we again let $K_1^L = K_1^E + 1$, comparing their wage level at τ,

$$ln(w_\tau^L) - ln(w_\tau^E) = \theta^L K_1^L - \theta^E K_1^E = \theta^L + K_1^E(\theta^L - \theta^E) = \theta^L + K_1^E(g_1^L - g_1^E),$$

we see that the difference is equal to θ^L (the return for the higher-ability/later mother), *plus* the difference in their wage level that has already emerged by $t = K_1^E$. Thus under this scenario, a regression of $l(w_{\tau i})$ on K_{1i} will provide a coefficient that not only over-estimates the average return to delay, $\bar{\theta}$, but even over-states the return for higher-ability mothers, θ^L.

Scenario #3: Unobserved Heterogeneity

Now consider a third scenario, in which the observed correlation between first-birth timing and wages is biased by unobserved heterogeneity. For simplicity, assume the extreme case in which fertility timing has no influence on wages: $\theta = 0$, namely $g_1 = g_2 \equiv g$. Under these conditions, the marginal benefit of delay is zero, and, per Equation (3), variation in K_1^* again arises only from variation in taste for early motherhood, ψ.

Yet to satisfy the universal observation that long-run wages and first-birth timing are positively correlated, in this scenario it must be the case that ψ is correlated with factors that influence women's wages. For instance, suppose ψ_i and a_i are negatively correlated,

[12] Although g_2 may vary with ability, in order for θ_i to increase with ability, the slope on a_i for g_1 must exceed the slope for g_2. Alternatively, g_1 may be equal for all, but g_2 may be increasing in K_1. For instance, if there is a positive correlation between taste for early motherhood, ψ, and taste for time at home with one's children, g_2 may stall by more for early mothers due to their greater decrease in labor supply. Likewise, in Section 2.1 I suggested that changes in labor supply may vary with K_1 because earlier mothers face a lower opportunity cost of time at home. Under this condition, wages will move together through K_1^E, but the early mother will have a larger kink in her wage path at K_1 because $g_2^E < g_2^L$. The level difference in their wages by τ will therefore equal $\theta^L + (\tau - K_1^E)(g_2^L - g_2^E)$.

suggesting the observed relationship between timing and wages is driven by ability bias. If this holds, late mothers will again be higher-ability women, who in turn have greater wage growth, $g_i = g(a_i)$, $g' > 0$.[13]

The third panel of Figure 2 shows the path of wages under Scenario #3. We see that there is now no kink in the wage path at K_1 for either mother, and instead wages diverge immediately upon labor market entry because $g^L > g^E$. If we again calculate the level difference in wages at τ,

$$l(w^L_\tau) - l(w^E_\tau) = \tau g^L - \tau g^E = \tau(g^L - g^E),$$

we see that it is unrelated to K_1, suggesting that a regression of $l(w_{\tau i})$ on K_{1i} should find no slope. Yet because of the correlation between ψ and ability – and because ψ drives K^*_1 – if we cannot fully control for ability, OLS will suffer from omitted variables bias, providing a positive coefficient on K_1.

<u>Scenario #4: Reverse Causality</u>

Lastly, consider a fourth scenario of reverse causality, in which the wage path drives fertility timing, rather than the reverse (Miller 2011 considers the milder case of simultaneity).[14] For instance, suppose that all women enter the labor force with a wage growth rate of g'_1, but that at some variable time in their career, each reaches a critical point, t_c, where their wage growth stalls to a slower rate g'_2. Suppose also that women choose to delay first birth until after t_c, thus late mothers are those who hit t_c later in their career.

The last panel of Figure 2 shows the path of wages under this case of reverse causality. As with Scenario #1, we see that the early-career wages of the two mothers are identical. Yet if women cannot perfectly anticipate t_c, in Scenario #4 their wages diverge *before* K^E_1. If we let $\vartheta \equiv g'_1 - g'_2$, the level wage difference at τ will equal $\vartheta(t^L_c - t^E_c)$, which again is

[13] Alternatively, as in Blackburn *et al.* (1993), women with high ψ may choose to invest less in human capital, thus $g_i = g(\psi_i)$, $g' < 0$.

[14] See, for instance, Heckman and Walker (1990) and Perry (2005) for evidence on wages influencing timing.

unrelated to K_1. But if women take roughly the same amount of time to conceive and give birth, $t_c^L - t_c^E \approx K_1^L - K_1^E$, a regression of $l(w_{\tau i})$ on K_{1i} will produce an OLS estimate consistent with ϑ, but that does not reflect a causal effect of fertility timing on wages.

3.3 Implications of the Potential Endogeneity of t_1

Lastly, I return to the question of the implications of the potential endogeneity of the timing of labor market entry, t_1. In my utility maximization model describing the choice of K_1^*, I simplified the question in two key ways. First, I started from a population for whom the optimal year of first birth, B_1^*, is after t_1, thus ignoring that initial step of the maximization decision. And second, I ignored any additional possible endogeneity of t_1.

In Section 2, however, I make the argument that the appropriate way to measure "timing" is in terms of labor market entry, using t_1 as an anchor to translate B_1 into "relative timing," K_1. Yet women have a joint choice over B_1 and t_1, and the *combination* of these two define K_1.[15] Furthermore, this choice sorts women into the two populations of mothers, "baby first" ($B_1 < t_1$, thus $K_1 < 0$) or "work first" ($B_1 > t_1$, thus $K_1 > 0$). Thus the potential endogeneity of t_1 is relevant to the interpretation of both those results that compare pre- and post-t_1 mothers, and the estimated returns to delay *within* each population.

In my analysis I will consider the question of how women sort into a pre- or post-t_1 first birth by assessing whether the choice of B_1^*, and B_1^* relative to t_1^*, appears to be more strongly a function of those individual characteristics that more directly influence the

[15] By comparison to the simple utility function in Section 3.1, to consider how women jointly choose B_1^* and t_1^*, one would need to consider a much richer model of lifetime utility. For instance, one would want to incorporate a broader range of earnings sources, since otherwise (barring perfect capital markets), no woman could choose $B_1^* < t_1^*$ because she would have no earnings or savings to live on. (In practice, approximately 50 percent of NLSY79 pre-t_1 mothers are living with a parent at first birth, with parental income their primary source of support; during the following pre-t_1 years support from spouses increases, with welfare reflecting a relatively modest source of income.) Such a model would also ideally incorporate marriage, with respect to both match quality, partner earnings, and the marriage-market consequences of a pre-marital first birth. (See, for instance, Caucutt *et al.* 2002 for a model that focuses on the marriage-market returns to fertility delay.) The question of marriage may be especially important, in light of my results suggesting that many women would have higher long-run earnings with a pre-t_1 first birth. Since we observe many women with a first birth shortly after t_1, this suggests that some *other* element of the utility function is driving this delay, such as the penalty for a pre-marital first birth.

marginal benefit of delay (such as ability), or those that influence the marginal cost of delay (*e.g.*, the number of children wanted).[16] I will also consider which of these factors are more strongly related to the length of time between finishing school and entering the labor market. In particular, note that the *magnitude* of K_1 is also defined by the endogenous choice of t_1. This is most obvious among pre-t_1 mothers, where, for a given year of birth B_1, K_1 becomes increasingly more negative as women wait longer to enter the labor market. If women who face higher potential wages start working more promptly, the endogeneity of t_1 may therefore influence my estimates of the slope parameters $g_0^{'}$ and θ (see Figure 1a).

4 Data and Variable Definitions

In this study I rely on data for the women from the NLSY79, who by 2008 had reached the ages of 44 to 51.[17] In the following section I first discuss how I calculate their relative timing of first birth, K_1, based on the year of labor market entry, t_1. I then discuss my dependent variable, w_{20}, the wage level approximately 20 years after labor market entry, and the selection issues surrounding the women for whom I lack an observed value. Last, I discuss my sample selection criteria, and report some summary statistics for my final sample.

[16] One might also worry that this sorting is correlated with the parameters of the wage equation themselves. For instance, suppose $g_{2i} \neq g_{0i}$, where the former is the post-motherhood wage growth rate a given woman i would experience if she had her first birth *after* t_1, and the latter is the element of her potential wage growth rate, $g_{pt_1 i}$, that is not a function of K_1. Furthermore, if the difference between these two parameters varies across the population, then sorting into a pre- or post-t_1 first birth may be endogenous to their relative size, such that $E[g_{0i}|pt_1 = 1] > E[g_{0i}]$ and $E[g_{2i}|pt_1 = 0] > E[g_{2i}]$. Since in my specification in Equation (4), one of my goals is to estimate δ, a function of $g_0 - g_2$ (depicted visually in Figure 1c as the discrete jump in $l(w_\tau)$ at $K_1 = 0$), since my OLS coefficient will reflect $E[g_{0i}|pt_1 = 1] - E[g_{2i}|pt_1 = 0]$, the *net* direction of the bias this sorting may generate is ambiguous. Alternatively, if $g_{0i} = g_{2i} \equiv g_i^{post}$ $\forall$ i, since my results instead suggest $\delta > 0$, my estimates must be biased by women with higher g_i^{post} systematically choosing $B_1 < t_1$: $E[g_i^{post}|pt_1 = 1] > E[g_i^{post}|pt_1 = 0]$. (These two possibilities are the equivalent of Scenarios #2 and #3 discussed in Section 3.2, where the estimate of δ is biased by endogeneity or unobserved heterogeneity, respectively.) Although I cannot directly test whether this sorting is correlated with these parameters, observable characteristics such as ability may drive much of their variation.

[17] Because the NLSY79 is well known, I leave a more detailed discussion of the data to the appendix. As discussed there, my sample builds from both the cross-sectional and minority supplemental samples, thus throughout I report weighted results. The appendix also provides a more detailed discussion of how I define the variables introduced here, and explores whether my sample selection criteria may influence my results.

4.1 Defining Relative Timing of First Birth, K_1

To determine the year a woman entered the labor market, t_1, I first establish her "graduation year," roughly the year in which she completed continuous schooling.[18] Note, however, that this raises the question of whether education and first-birth timing are jointly determined. For instance, if some women get unexpectedly pregnant while in school, and that pregnancy derails their education, then a woman's education level at t_1 may be endogenous for those pre-t_1 mothers. Education may also be endogenous to *intended* first-birth timing: women who want an early birth may choose less schooling (Blackburn *et al.*, 1993).

As I show in Appendix Section F, however, in the NLSY79 I find little evidence of either potential source of endogenous schooling. For instance, consistent with Rindfuss *et al.* (1980) and Stange (2011), among pre-t_1 mothers, surprisingly few were still in school when they conceived, and of those, many subsequently completed more schooling before t_1. These results are consistent with the finding of at most weak negative effects of a teen birth on education (see, for instance, Geronimus and Korenman 1992, and Hotz *et al.* 2005).[19]

After defining a woman's graduation year, from that point forward, I search for the first twelve-month period in which she worked at least 1000 hours, defining this as t_1.[20] Given t_1, and the calendar year of a woman's first birth, B_1, I then define K_1, the relative timing: $K_1 = B_1 - t_1 + 1$. Figure 3 shows the distribution of K_1 for the population of NLSY79 women observed through age 40.[21] Note the dip in the distribution at $K_1 = \{0, 1\}$ – relatively few

[18]In determining "graduation year," I allow gaps in post-high school education, as long as upon return, a woman completes more years of schooling at a sufficient pace (see the appendix for more detail). For example, for a woman who takes a year off after high school, then returns to complete a college degree, her "graduation year" will be based on the latter, as long as she was not working full-time throughout college.

[19]To test for the second potential source of endogenous schooling, I consider whether women who may prefer an earlier first birth also intend to get less schooling. I find that among "high school" or "some college" types (but not "college" types), those with a higher taste for early motherhood may systematically choose less schooling. Although throughout this analysis I group women with 12 to 15 years of schooling at t_1, this endogeneity may mean that some early mothers are grouped into an education level below their 'potential,' and may therefore have higher wage growth than the average woman in their observed education group.

[20]Using this criterion, of the NLSY79 women observed till age 40, only 106 never enter the labor force. If I limit the analysis to women who work at least 1500 hours in t_1, my results are substantively equivalent.

[21]Figure 3 uses the NLSY79 cross-section sample. At one extreme I show the proportion of women who never work (NW), and at the other the proportion who never have kids (NK). For the latter, the first bar reflects those who have ever married by their 40s, and the second, those who have not married by this point.

women enter the labor force for the first time with a newborn or when pregnant. This shows that t_1 is endogenous to B_1, at least within this short range.

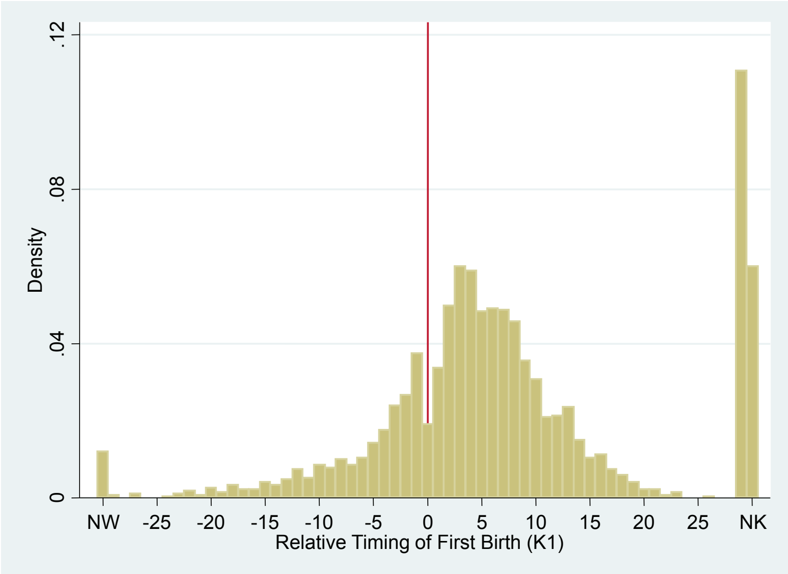

Figure 3: Distribution of Relative Timing of First Birth, K_1

4.2 Defining Long-Run Wages, w_{20}

My dependent variable is a woman's hourly wage approximately 20 years after labor market entry, w_{20}. Specifically, I use a wage observed in a six-year window between a woman's 19^{th} and 24^{th} "career years" (calendar years measured from t_1), defining this as "t_{20}."[22] Although using this six-year window helps to limit the bias that may arise from capturing only women with strong labor market attachment, as we see in Table 1, the proportion of women captured in my sample still varies appreciably by timing of first birth.

In particular, I am still restricted to those women (i) who I observe at least 19 years beyond t_1, and (ii) who work at some point during this six-year stretch. The first issue has the strongest implications for pre-t_1 mothers, where I systematically lose the 7 percent who take longer than 15 years to enter the labor force after B_1. Furthermore, as we see from Table 1, excluded women among the earliest mothers have significantly lower AFQT

[22]Starting from t_{21}/t_{22} and working outwards (these data come from the period after 1993, when the survey became biennial), I use the first observed wage within this 6-year range as w_{20}. See footnote 48 for a discussion of my results if I instead use wages observed 15 years after t_1.

Table 1: Comparison of Women Included and Excluded from the Analysis Sample

Pre-t_1 First Birth:	$K_1 \leq^{-} 8$			$K_1 =^{-} 7 -^{-} 4$			$K_1 =^{-} 3 - 0$		
	In		**Out**	**In**		**Out**	**In**		**Out**
AFQT score	30.0	**	22.6	32.3		30.4	37.8		31.9
Expected schooling ('79)	12.2		12.5	13.1		12.9	13.3		13.3
Mother's education	10.1		9.9	10.8		10.6	10.7		11.0
Mother's age at K_1	19.7		20.4	20.5		21.6	20.1		19.8
Minority (%)	38.1	**	50.4	46.0		37.1	39.3		37.4
Education at t_1	11.5		11.7	12.0		11.9	12.0		11.8
Starting wage ($2000)	6.76		7.09	7.24		7.95	7.18		7.23
Relative timing (K_1)	-9.4	***	-14.4	-5.2		-5.4	-1.4	*	-1.7
Age at K_1	18.0	***	19.4	18.7	***	19.9	18.9	***	20.6
Married at K_1 (%)	56.4	**	40.7	52.0		49.5	54.9		60.7
Husband's earnings (t_{20})	21.0		24.1	21.7		26.2	26.8		22.9
Children by t_{24}	3.1		3.1	2.6		2.6	2.5	*	2.8
Yrs obs. beyond K_1	30.9	***	27.5	28.9	***	26.2	28.4	***	26.3
Sample Size:	**110**		**387**	**214**		**80**	**485**		**87**
(% included)		(25.1)			(70.7)			(85.1)	

Post-t_1 First Birth:	$K_1 = 1 - 3$			$K_1 = 4 - 7$			$K_1 = 8 - 11$			$K_1 \geq 12$		
	In		**Out**	**In**		**Out**	**In**		**Out**	**In**		**Out**
AFQT score	47.3		51.8	56.1		60.5	57.4		60.6	56.0		59.2
Expected schooling ('79)	13.9		14.2	14.1	**	14.6	14.3	**	15.0	14.1	**	14.6
Mother's education	11.4		11.6	11.9		12.3	12.0	**	12.8	12.2		12.1
Mother's age at K_1	21.0		21.1	21.8	***	22.9	22.3	*	23.4	22.4		22.7
Minority (%)	21.7		19.4	12.9		13.0	9.9		11.2	11.0		12.3
Education at t_1	12.8		12.9	13.3	**	13.8	13.3	***	14.0	12.8	***	13.6
Starting wage ($2000)	8.43		8.16	8.90		9.49	8.63	*	9.40	7.90	***	9.22
Relative timing (K_1)	2.2		2.3	5.4		5.4	9.2		9.3	13.5	***	16.3
Predicted timing (K_1^{pred})	3.9		3.7	4.9	**	4.0	5.6		5.3	5.6		5.6
Age at K_1	21.9	**	22.9	25.3	***	26.9	29.0	***	30.1	32.6	***	36.2
Married at K_1 (%)	79.4		78.0	87.6		88.3	86.3		89.9	85.0		83.7
Husband's earnings (t_{20})	33.2	***	63.0	45.4	***	72.5	42.6	***	75.6	44.7	**	60.0
Children by t_{24}	2.5		2.8	2.2	**	2.5	2.0	**	2.2	1.5	*	1.7
Yrs obs. beyond K_1	25.3	**	23.5	21.9	***	19.4	18.1	***	15.9	14.7	***	10.2
Sample Size:	**529**		**81**	**638**		**122**	**368**		**92**	**189**		**189**
(% included)		(87.0)			(83.7)			(81.2)			(51.5)	

NOTES: This table compares mean characteristics, by first-birth timing, between those women included in my sample and those excluded (these samples reflect the exclusion restrictions discussed in Section 4.3). My comparison starts from the subset of NLSY79 women observed through at least age 40, and who have had a child by their last year observed; I group women who never work into the earliest timing group. All values reflect weighted means; see the appendix for more detail on these weights and how I build these variables. AFQT scores reflect age-adjusted percentiles (measured in 1979). "Husband's earnings (t_{20})" reflects the average, over the period from t_{19} to t_{24} of either spouses's earnings (if married) or child support/alimony (if not). "Children by t_{24}" reflects total children born by t_{24} (or a woman's last year in the survey). Predicted timing of first birth (for post-t_1 mothers), is as predicted in the year of labor market entry. Significance levels reflect whether the given variable is significantly different between those include and excluded from the sample (indicated by * at the 10% level, ** at 5%, and *** at 1%).

scores, and are more likely to be minorities. Yet as I discuss in more detail in Appendix Section E, controlling for background characteristics, the excluded women show no evidence of systematically different wage patterns, either when they start working or by 10 years later, suggesting that their wages by t_{20} will also be similar. Thus their exclusion likely has little influence on my results.

The second issue – mothers who do not work during "t_{20}" – is more relevant among post-t_1 mothers. As one would anticipate, we see that excluded women tend to have more children and greater spousal financial support. We also see, however, that they have more education and higher starting wages.

Yet these latter differences are primarily driven by the educational distribution; if I recreate Table 1 separately for the three education levels used in my analysis, these differences largely disappear.[23] (By contrast, the difference in total children and spousal earnings are evident within education levels as well.) Thus, as I discuss in the appendix, grouping by education, I find that women with no observed w_{20} have similar wage patterns over time (*e.g.*, starting wages or wage growth to t_5 if $K_1 > t_5$), suggesting that their exclusion likewise will not influence my results.

4.3 Final Sample Selection

My final sample starts with the set of NLSY79 mothers for whom I can observe w_{20}. I then apply two additional sample restrictions: (*i*) I exclude women with a first birth beyond their 17^{th} career year, and (*ii*) I exclude women who are over 35 when they have their first child. I apply the first restriction, which drops only 2 percent of mothers, because among

[23] The differences in age at first birth are largely mechanical (and remain when I separate the sample by education level); because my data are for a single birth cohort, women who are older at first birth are more likely to "age out" of the sample before their 19^{th} (or 24^{th}) career year. (This in turn speaks to the differences in education: *within* each timing group, women who are older at first birth must have entered the labor market later, and the primary reason they did so is because they completed more schooling before t_1.) For the last timing group ($K_1 \geq 12$), however, I still see the same pattern of sorting within high school graduates (but not within college graduates). The difference is driven by my exclusion restriction of mothers who were over 35 at first birth (discussed in Section 4.3), which disproportionately affects women who had some college education at t_1, who in turn had more education and higher wages at t_1.

late mothers, I do not want to capture a sample that is influenced by the rate at which women return to work after motherhood. For instance, if, among later mothers, those with higher potential wages return more quickly, this may generate a spurious positive correlation between K_1 and w_{20}.

I apply the second restriction because I find evidence suggesting that among women from the NLSY79 cohort, those with a first birth beyond age 35 have some underlying characteristic which led them to both have a late first birth, and have lower wage growth.[24] In particular, evidence suggests that as women approached the end of their 30s and the oncoming fertility decline (hence my focus now on *age* at first birth), some chose to have a child under lesser circumstances. As shown in Appendix Figure A-1, I find a sharp drop after age 35 in the proportion of women who were married when they conceived their first child or at the birth of that child, and among those married, the older mothers had relatively lower-earning spouses (this pattern holds at all education levels).

Furthermore, the data suggest that these older mothers possess some characteristic that was penalized in *both* the marriage and labor markets (consider the literature on the relationship between earnings and beauty, *e.g.*, Hamermesh and Biddle, 1994). Specifically, among college graduates, Appendix Figure A-2 shows that women who had a first birth after 35 already suffered discontinuously lower wage growth long before motherhood. This leads to the observation that these later mothers have lower long-run wages, but not *because* of their late fertility timing. I therefore exclude these women from my sample, because their inclusion would lead to spuriously low estimates of the return to fertility delay.

These restrictions provide a sample of 2,567 mothers, of which 815 (25 percent weighted) have their first child before t_1. For this sample, Table 2 reports mean values of w_{20} and starting wages, as well as age at first birth, age at t_1, and several background characteristics,

[24] Among all mothers observed through age 40, 4.7 percent have a first child after age 35; 2.0 percent of high school dropouts, 4.0 percent of high school graduates, and 8.7 percent of college graduates. (By comparison, the proportion with a first birth after t_{17} is very even across education groups.) This selection pattern may no longer hold for more recent cohorts, in which first births after age 35 are more common.

Table 2: Summary Statistics by Timing of First Birth

	All	**Fertility Timing (K_1)**						
		$\leq {}^{-}7$	${}^{-}7$-${}^{-}4$	${}^{-}3$-0	1-3	4-6	7-9	10+
High School Dropouts:								
$\ln(w_{20})$	2.25	2.13	2.26	2.32	2.15	2.36	2.30	2.22
w_{20} (2000$)	10.70	9.02	10.32	11.68	9.43	12.16	10.76	10.51
w_0 (2000$)	6.75	6.93	6.68	6.56	7.13	7.02	6.11	6.16
$\ln(w_{20})$-$\ln(w_0)$	0.41	0.27	0.42	0.50	0.24	0.49	0.56	0.37
Total education at t_1	10.5	10.4	10.7	10.5	10.3	10.6	10.0	10.7
AFQT score (percentile)	29.9	22.8	26.2	31.3	30.6	32.7	28.6	39.1
Highest grade expected	12.0	11.6	12.1	11.9	12.3	12.2	11.9	12.5
Mother's education	10.3	9.7	10.3	10.2	10.5	10.5	9.5	10.9
Minority (%)	24.5	37.1	41.5	28.0	18.9	12.3	18.1	8.2
Age at first birth	20.4	17.4	18.0	17.9	20.1	22.8	25.4	29.5
Age at t_1	20.1	26.4	23.1	19.6	18.1	18.3	17.9	17.7
Sample Size:	**506**	**82**	**64**	**144**	**86**	**67**	**35**	**28**
% of all (weighted)		13.3	10.9	27.3	18.1	15.8	8.2	6.3
High School Graduates:								
$\ln(w_{20})$	2.46	2.34	2.30	2.45	2.41	2.42	2.51	2.58
w_{20} (2000$)	13.15	11.65	10.87	13.30	12.42	12.51	13.93	14.86
w_0 (2000$)	7.74	7.01	7.00	7.32	7.87	8.06	8.01	7.59
$\ln(w_{20})$-$\ln(w_0)$	0.47	0.52	0.45	0.53	0.40	0.39	0.52	0.59
Total education at t_1	12.5	12.4	12.5	12.4	12.5	12.5	12.5	12.5
AFQT score (percentile)	47.0	34.3	34.3	39.2	45.6	50.6	50.9	53.0
Highest grade expected	13.7	13.4	13.4	13.6	13.8	13.6	13.7	13.9
Mother's education	11.3	11.3	10.5	10.8	11.1	11.6	11.4	11.8
Minority (%)	22.2	42.3	45.4	43.2	22.9	15.3	12.4	12.3
Age at first birth	24.1	18.6	19.3	19.1	21.5	24.1	26.9	31.2
Age at t_1	20.0	27.8	24.3	20.7	19.5	19.4	19.3	19.0
Sample Size:	**1,731**	**55**	**112**	**333**	**396**	**326**	**261**	**248**
% of all (weighted)		2.4	4.7	15.0	22.6	21.3	17.1	16.9
College Graduates:								
$\ln(w_{20})$	2.86	2.76	2.53	2.53	2.78	2.77	2.95	3.07
w_{20} (2000$)	20.52	21.54	14.43	14.33	18.77	18.26	22.11	25.25
w_0 (2000$)	11.26	13.53	11.45	9.52	11.63	11.19	11.91	10.22
$\ln(w_{20})$-$\ln(w_0)$	0.51	0.26	0.17	0.36	0.41	0.41	0.53	0.83
Total education at t_1	16.2	17.0	16.0	16.2	16.3	16.1	16.2	16.2
AFQT score (percentile)	77.2	61.1	57.1	67.4	75.2	79.3	77.4	79.0
Highest grade expected	16.3	15.7	16.0	16.4	16.2	16.1	16.5	16.4
Mother's education	13.3	11.5	13.3	13.3	13.6	13.0	13.2	13.7
Minority (%)	8.6	28.3	58.3	13.9	11.7	7.0	6.5	7.1
Age at first birth	28.9	19.3	20.1	22.2	25.5	27.7	30.7	33.8
Age at t_1	23.0	28.5	24.7	23.8	23.4	22.8	22.9	22.5
Sample Size:	**330**	**5**	**6**	**14**	**61**	**98**	**91**	**55**
% of all (weighted)		1.1	0.9	4.1	16.7	30.8	29.2	17.3

NOTES: All values reflect weighted means. See the appendix for more detail on these weights and how I build these variables. AFQT scores reflect age-adjusted percentiles (measured in 1979), and highest grade expected is reported at approximately age 18. Monetary values are translated into year-2000 dollars using the Consumer Price Index for all urban consumers.

grouping women by their first-birth timing and their education level at t_1.[25] Throughout this analysis I use three education categories: less than a high school diploma or a GED (Cameron and Heckman, 1993), exactly a high school diploma or some college education, and at least a college degree. Appendix Table A-1 provides additional summary statistics.

Looking first at women who remain childless at t_1 ($K_1 > 0$), for both high school and college graduates Table 2 echoes the existing literature's finding that long-run wages are increasing in first-birth timing. Excepting high school graduates, we also see no clear link between K_1 and either AFQT scores or educational expectations, providing a first suggestion that the correlation between timing and wages is not driven by ability bias.

Among pre-t_1 mothers, Table 2 also shows a positive link between wages and timing among high school dropouts, and maybe also high school graduates. (Given the small number of college graduates with a pre-t_1 first birth, throughout the following sections I ignore this education level when discussing pre-t_1 mothers.) We also see, however, that AFQT scores are increasing with timing within these two groups, suggesting that the link with wages may only reflect this slope on ability.

Lastly, if we compare the outcomes for women with their first birth just before versus just after t_1, we see initial evidence suggesting that it may be relatively better to have one's first child before entering the labor market. In particular, among both high school dropouts and graduates, women with their first birth within 3 years before t_1 have significantly higher long-run wages than women with their first birth in the 3 years following.

5 Identification Strategy

In the following section I first discuss means to consider the implications of the possible endogeneity of t_1. I then discuss my identification strategy for assessing the possible bias captured in the estimates of the return to delay for pre- and post-t_1 mothers.

[25] See also Figures 4 and 6 for plots of various of these characteristics by K_1.

5.1 Checking for the Possible Endogeneity of t_1

To consider the implications of the possible joint determination of B_1 and t_1, I first test the relationship between individual characteristics and the following three outcomes:

1. the length of time between finishing school and having one's first birth (a linear transformation of B_1);
2. among women with high taste for early motherhood (as suggested by an observed first birth within 4 years of graduation[26]), the probability that $B_1 < t_1$; and
3. the length of time between finishing school and entering the labor force ("years off"), separately among pre- and post-t_1 mothers.

I consider this among the 1,134 NLSY79 women who complete their education at high school graduation (before they turn 20), and who have not conceived their first child by that point. I use high school graduates both because this education level has greater balance between pre- and post-t_1 mothers, and because this is a standard educational stopping point, suggesting that fewer of these women will be finishing school *in order* to have their first child. (Yet as reported in the notes to Table 3, the results are very similar if I instead run these regressions on my analysis sample.)

The first two comparisons provide insight into the extent to which variation in B_1, and the choice of $B_1 < t_1$, arises primarily from variation in factors that more directly influence the marginal cost of delaying one's first birth, or the marginal benefit. And, given the sorting into a pre- or post-t_1 first birth, the last comparison considers the extent to which the pace of subsequent entry into the labor force may be endogenous to a woman's potential wage. Using my analysis sample of high school graduates, I also plot several of these observable characteristics by K_1, to provide insight into the pattern of sorting both between, and within, these two populations of mothers.

5.2 Checking for Bias in the Estimated Slopes on K_1

Now I return to the question of how to identify the effect of first-birth timing, given women's capacity to control their fertility. The existing literature has relied on various identification

[26] As reported in the notes to Table 3, the results are very similar among those who *predict* a first birth within 4 years.

strategies to tackle this issue. Some research has used a fixed effects approach, as in Geronimus and Korenman (1992), Taniguchi (1999), and Amuedo-Dorantes and Kimmel (2005). Other papers have used instrumental variables, for instance relying on health shocks such as miscarriages, as in Hotz *et al.* (2005) and Miller (2011).

Yet, as nicely explained in Wilde *et al.* (2010), this literature suffers from the lack of strong instruments that influence fertility timing without having a direct effect on women's economic outcomes.[27] In my analysis, I therefore follow a simple ordinary least squares (OLS) approach, relying largely on the richness of the NLSY79 data to capture much of the underlying variation between early and late mothers. Although this cannot rule out that my results may be influenced by systematic differences in *unobservable* characteristics, to drive my results such factors must be uncorrelated with all of the observable characteristics that I can account for. Furthermore, my conclusions are based not only on the correlation between first-birth timing and long-run wages, but also on the pattern of wages over time.[28]

In particular, as a first gauge of the relationship between first-birth timing and wages, I plot the path of wages for the women from the NLSY79, grouping them by their relative timing of first birth, K_1. For the post-t_1 mothers, to assess whether the positive correlation between timing and wages may be contaminated by bias, I then compare this figure to the hypothetical wage paths in Figure 2, each associated with a potential source of bias discussed in Section 3. Because two of these scenarios also have implications for the relationship between timing and ability, I also plot average AFQT scores by K_1.

Building on the model from Section 2.2, I then estimate the following wage equation:

$$l(w_{20}) = l(w_0) + \theta k_1 + \gamma k_1^{pt_1} + \delta pt_1 + X\beta, \tag{4}$$

[27]They also highlight the limitations of using non-mothers as a comparison group, as is evident in the NLSY79 as well. For instance, if I include childless women in Figure 5a, their wages start below those for women with $K_1 = 5 - 10$, continuing to fall behind over the next five to ten years, catching up only after each group of mothers have their first child.

[28]In practice, because my main specifications also control for starting wage, my method might be considered a fixed-effects approach. In addition, footnote 45 reports the results if I estimate θ following the instrumental variables approach introduced in Miller (2011), as amended in Herr (2008).

where $k_1 \equiv K_1(1 - pt_1)$, $k_1^{pt_1} \equiv K_1 pt_1$, $\gamma \equiv -g_0'$, and $\delta \equiv (g_0 - g_2)t_{20}$. Referring back to Figure 1, γ and θ reflect the slope on K_1 for pre- and post-t_1 mothers, respectively (where I replace $-g_0'$ with γ to limit confusion arising from the negative sign), and δ reflects the level difference in long-run wages for women with their first birth just before versus after t_1 (the discontinuity at $K_1 = 0$ in Figure 1c).

As the second test for bias in the estimates of the slope parameters, I use the following OLS approach to consider how the estimated relationship between timing and wages shifts as I control for an increasing number of individual characteristics. My attention throughout this exercise is on changes in the coefficient estimates of γ and θ, as evidence to whether the initial "raw" coefficients are capturing bias as well as, or instead of, a causal effect.

After first reporting the timing coefficients estimated with no controls at all, I estimate Equation (4) including very few additional controls beyond those that speak to the distribution of wages from which a woman's values of w_0 and w_{20} were drawn, as well as several family characteristics including spousal information.[29] I treat these OLS-estimated coefficients for γ and θ as my "raw" measures of the partial correlation between the timing of first birth and long-run wages.

As a first check on the robustness of these coefficients, I then control for a set of observable characteristics, X_{a_i}, that reflect ability, career motivation, and other underlying characteristics that may directly influence a woman's wage path. These include:

1. her age-adjusted AFQT score (measured in 1979), as a proxy for ability;
2. her educational and career expectations, measured as expected total schooling (reported at approximately age 18), and expected labor force participation and occupation at age 35 (reported at approximately t_1);

[29] The former include where she lived at t_1 and t_{20}, and the calendar and "career" year of t_{20}. The latter includes total children by t_{20} (so that my coefficients do not capture the indirect effect on wages of timing's influence on total fertility), and two measures of spousal characteristics. The first is the average of "spousal support" (his earnings when married, and alimony or child support when not), for the years from K_1 through t_{20}, when the availability of such income may especially influence a woman's labor supply decision. The second factor, reflecting spousal "quality," captures the fact that via assortative mating, "higher income-type" men tend to marry "higher income-type" women. Lastly, in all specifications I also control for years education at t_1 (and, for high school dropouts, a GED). See the notes to Table 4 for additional controls included in these regressions, and Appendix Section C for more detail on how I define them.

3. the length of time between her "graduation year" and t_1, "years off," as a reflection of motivation or other characteristics of her underlying labor market quality; and
4. a measure of her self-esteem and her score on the Rotter locus of control scale (measured in 1980 and 1979, respectively), where the latter reflects the degree to which she believes she has control over the events of her life.

If the wages of 'early' and 'late' mothers diverge due to heterogeneity (as in Scenario #3 of Figure 2), the elements of X_{a_i} may be some of the key factors driving the separation of wages over time. Furthermore, X_{a_i} may influence the optimal timing of a woman's first birth if they enter into her marginal benefit of delay, leading to endogenous sorting as in Scenario #2. Thus if I find that the OLS-estimates of γ or θ fall with the inclusion of X_{a_i}, this suggests that the raw estimates were biased upwards by either endogeneity or omitted variables bias.

I next control for a set of characteristics, X_{ψ_i}, that either indirectly reflect taste for early motherhood, ψ_i, or otherwise influence K_{1i}^* through the marginal cost of delay. These include how many children a woman expects to have, her mother's age at first birth (if ψ is inherited), her views on gender–role norms, and her age at first intercourse. If, as in Scenario #3, the positive coefficient on timing arises from a correlation between ψ and factors that influence wages, including X_{ψ_i} should drive the OLS-estimates of γ and θ further towards zero by absorbing more of this heterogeneity.

Benefiting from the richness of the NLSY79, I next control for a broader set of background characteristics, X_{b_i}, that may also indirectly build into a woman's K_1^* by generating variation in either the marginal benefit or marginal cost of delay. These include her family background, such as her race and ethnicity, the religion in which she was raised, and her mother's education and labor force participation.

Notice that as I control for an increasing number of factors that influence K_1^*, the OLS-estimated coefficients on timing will increasingly be identified off of the difference between a woman's observed and optimal timing of first birth. As a last step, I take this further to control directly for a woman's predicted timing, K_1^{pred}.[30] Taking this value as reported at t_1 or the year before, for a woman with $K_1 > 0$, this reflects her estimate of K_{1i}^*, which will be driven in part by her expected return to delay, $E_{t_1}[\theta_i]$.

Lastly, note that throughout this progression I never control for factors that may reflect part of the *mechanism* of how first-birth timing affects the wage path (Buckles, 2008), because doing so may absorb part of the full effect of timing on a woman's long-run wage level. For instance, because timing may influence labor supply choices, I do not control for accumulated work experience, or elements such as tenure or part-time status at t_{20}.

Thus to summarize, my focus throughout this process is on how the estimates $\hat{\gamma}$ and $\hat{\theta}$ shift with the inclusion of this series of controls. To the extent that they change, the stage at which this occurs will provide insight into which types of bias are captured in the raw coefficients. But if the coefficients remain stable, despite the addition of these rich measures of individual characteristics, this suggests that the original estimates largely reflect only a causal effect of the timing of first birth.

6 Results

In the following section, I first discuss the evidence on the possible bias captured in my results. In Section 6.2 I then discuss my conclusion on the appropriate way to measure timing of first birth, and in Section 6.3 I discuss the economic implications of my findings.

6.1 Testing for Bias

Evidence and Implications of the Possible Endogeneity of t_1

For the sample of NLYS79 women who completed their schooling at high school graduation and had not yet conceived their first child, Column (1) of Table 3 reports the coefficients from

[30] In most survey waves, women were asked when they expected to have their next child, which for childless women reflects their predicted timing of first birth. To the extent that I am now estimating $\hat{\theta}$ using only the unexpected-at-t_1 factors that create the difference between predicted and observed timing, one concern is that these unexpected elements may be correlated with a woman's individual return to delay, θ_i. For instance, women may learn new information about θ_i after entering the labor market, and may adjust their timing accordingly. Yet I find no evidence to this effect. For instance, among college graduates with $K_1 > 3$, wage growth to t_3 is unrelated to the difference between K_1^{pred} and K_1.

Table 3: Evidence on the Endogeneity of t_1

Dependent: Variable (Y):	**Years till baby** (1)	**"Baby First"** ($pt_1 = 1$) (2)	**Years Off if** $pt_1 = 1$ (3)	**Years Off if** $pt_1 = 0$ (4)
Mean of Y	6.9	31.1%	7.1 yrs	0.4 yrs
AFQT score (x 10^1)	0.051	-0.024**	-0.751***	-0.039**
	(0.079)	(0.012)	(0.264)	(0.016)
Unemployment rate in	0.002	0.001	0.199	0.046***
graduation year	(0.057)	(0.008)	(0.146)	(0.014)
Age at first sexual	0.284***	-0.037**	-0.390	-0.005
intercourse	(0.083)	(0.016)	(0.391)	(0.018)
Mother's age at first birth	0.100**	-	0.137	0.007
	(0.047)	-	(0.126)	(0.011)
Mother age 17 to 19	-	0.096*	-	-
at first birth	-	(0.049)	-	-
Mother worked FT in 1978	0.075	-0.085*	-1.621	-0.121*
	(0.338)	(0.049)	(1.328)	(0.070)
Gender/family role attitudes	-0.088*	0.003	-0.048	-0.019
	(0.047)	(0.007)	(0.091)	(0.014)
Raised in religious household	-0.886	-0.215**	3.393***	-0.180
	(0.573)	(0.101)	(1.189)	(0.136)
Minority	-0.709	0.065	-3.623**	-0.023
	(0.581)	(0.085)	(1.534)	(0.162)
Sample Size:	**1134**	**519**	**197**	**937**

NOTES: Each column reports the OLS coefficients (and robust standard errors) from a weighted regression of the given dependent variable on the controls listed above, as well as a woman's age and location at graduation (region and SMSA), her father's occupation, mother's education, family characteristics when she was 14, and her Rotter score and self esteem. (A higher gender-roles score reflects more conservative views.) Column (1) uses the subset of all NLSY79 women observed through age 40 who complete their schooling at exactly high school graduation (before they turn 20), and who have not yet conceived their first child. The Column (2) sample are the subset of these women who have their first child within 4 years of graduation, to focus on women with relatively high taste for early motherhood. The results are similar if I instead use women who *predict* a first birth within 4 years (for the subset for whom I can observe such predictions at graduation); if I instead use the full Column (1) sample, the coefficients fall, but the same characteristics are significantly related to the probability of a pre-t_1 first birth (excepting religion). The Column (3) and (4) samples are the subset of the Column (1) sample who have a pre- and post-t_1 first birth, respectively. If I rerun these specifications on my analysis sample of high school graduates and dropouts, the results are very similar. For instance, in Columns (1) and (4), the same factors are significantly related to the given dependent variable; in Columns (2) and (3) the same holds, except for mother's teen birth or full-time working (in the former), and religious background (in the latter), where the signs remain the same. For Columns (2) through (4), however, I find a weaker link between AFQT scores and the given dependent variable within my analysis sample (by 25 to 60 percent), as well as a weaker link with the unemployment rate in Column (4). Rerunning Column (1) on college graduates, I find only a significant link with age at first sex. Likewise, in Column (4) I find only evidence of a negative link with mother's labor force participation and a positive link with religious background (each at the 15-percent level); I find no link with either AFQT scores or the unemployment rate. (There are too few pre-t_1 college graduates to run either Column (2) or (3).) Significance levels are indicated by * (at the 10% level), ** (at 5%), and *** (at 1%).

an OLS regression of the number of years till first birth, a linear transformation of B_1. (See the notes to Table 3 for a discussion of these results if run instead on my analysis sample.) Among these women, we see evidence suggesting that variation in B_1 arises primarily from variation in factors that more directly influence the marginal cost of fertility delay, rather than the marginal benefit.[31] For instance, we see a strong relationship between timing and one's views on appropriate gender and family roles, mother's age at first birth, and age at first intercourse, but no link with AFQT scores.

For the subset of these women with high taste for early motherhood – as suggested by an early first birth – Column (2) reports the coefficients from an OLS regression of $pt_1 = 1$, reflecting the factors that influence whether these women choose "baby first" or "work first." We again see evidence suggesting that this outcome is largely driven by variation in taste. For instance, women from religious backgrounds are more likely to have a baby first, whereas women whose mothers worked are more likely to work first. And although there is no linear link with mothers' age at first birth, women whose mothers had a first birth in their late teens are likewise significantly more likely to have a first birth directly out of high school, before entering the labor force. Yet among this population, we do see that the propensity to choose "baby first" is decreasing in ability, suggesting that their timing of t_1 relative to B_1 may be in part a function of their potential wage.

Columns (3) and (4) next report the coefficients from OLS regressions of the number of years off between finishing school and entering the labor force, separately for pre- and post-t_1 mothers. As we see from the first line of each column, the average length of time off was 7.1 years among pre-t_1 mothers, but only 0.4 years for post-t_1 mothers, where 78 percent went straight from school to work.[32]

[31]Variation may also arise from factors outside of the simple model in Section 3, such as marriage timing, or spousal preferences and earnings. For instance, if I rerun these specifications controlling for marital status (by the end of the summer after graduation), and spouse's income that year, I find that women who marry promptly have their first birth significantly sooner, and are more likely to have a baby first rather than work first. Yet among pre-t_1 mothers, neither marital status nor husband's earnings affects years off, even if measured in the year of first birth. (Controlling for these additional factors does not substantively change the results reported in Table 3.) I do not include marital status in the main specifications because women may marry promptly *in order* to have an early first birth.

[32]Within my analysis sample, 80 percent of post-t_1 mothers go straight to work (54 percent of high school

For both populations, we see that the timing of entry into the labor force is strongly related to a woman's AFQT score, as well as the local unemployment rate.[33] This again suggests that t_1 is endogenous to a woman's potential wage, as we would expect, especially among pre-t_1 mothers, where the higher potential wage is necessary to offset the higher reservation wage associated with having small children at home. We also see, however, that the timing of t_1 is influenced by taste; pre-t_1 women raised in religious households start working later, and post-t_1 women whose mothers worked start working sooner.

Overall, the results in Table 3 suggest that both the sorting between "baby first" and "work first" (which defines the sign of K_1), and the subsequent timing of entry into the labor force (which defines its magnitude), are at least somewhat endogenous to a woman's potential wage. This raises the possibility that my use of t_1 to translate the year of first birth into relative timing, K_1, may introduce an additional source of bias into my estimates.

In Figure 4 I therefore use my analysis sample of high school graduates to plot the pattern of observable characteristics across the threshold at $K_1 = 0$, to consider whether this transformation of B_1 into K_1 has generated a discontinuity in the "type" of women with a first birth just before versus just after labor market entry.[34] (The patterns are equivalent for my analysis sample of high school dropouts).

For instance, the first panel plots average AFQT scores by relative timing, showing that ability rises with timing among both pre- and post-t_1 mothers, mirroring the results in Table 3. Yet there is no level discontinuity in AFQT scores at $K_1 = 0$. We see the same in the other three panels of Figure 4, which plot the average local unemployment rate in t_1,

dropouts, 81 percent of high school graduates, and 89 percent of college graduates).

[33]For pre-t_1 mothers, the link between years off and the unemployment rate is almost significant at standard testing thresholds. Note also that the magnitude of the effect of AFQT scores is very similar within the two columns, when measured as a proportion of the mean of years off.

[34]Each panel of Figure 4 plots the average of the given observable characteristic separately for women with each value of K_1 within 8 years to either side of t_1. I include self esteem based on results showing that it is strongly correlated with a woman's long-run wage level at this education level. (For highest grade expected at approximately 18, I use only those women who remain childless at that point, although the figure is equivalent using the whole sample.) In each panel, I include the fitted line on either side of t_1 and its 95-percent confidence interval. Only the slopes on AFQT scores are significantly different from zero (and, among high school dropouts, only for pre-t_1 mothers).

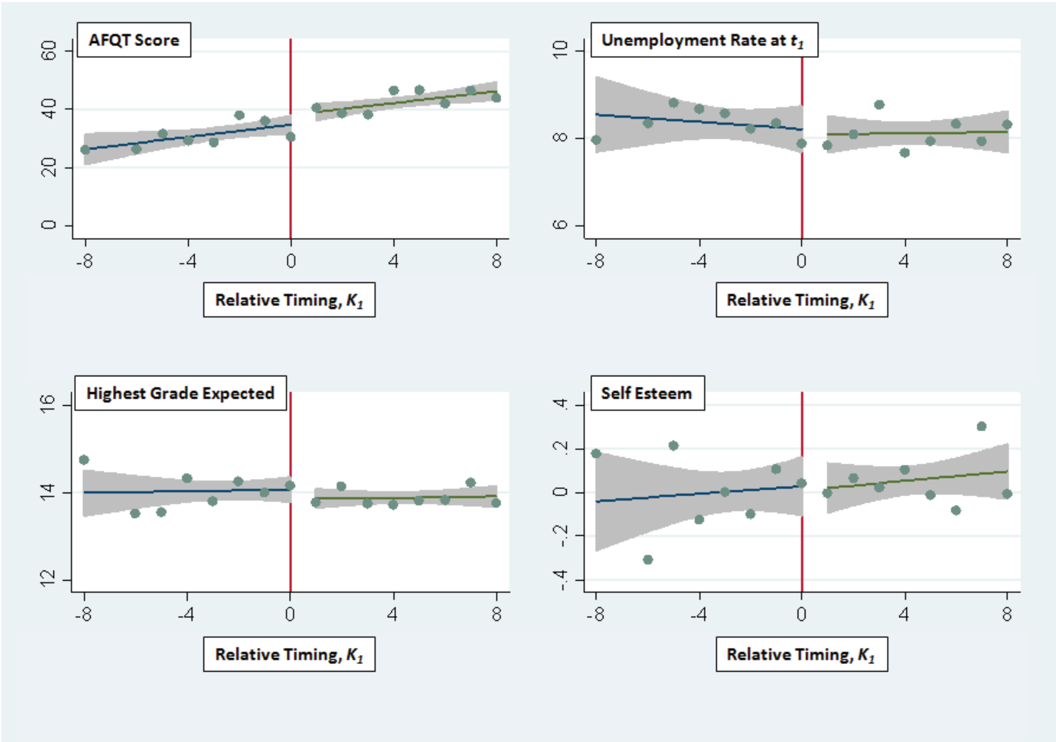

Figure 4: Average Characteristics by Relative Timing of First Birth

a woman's highest grade expected at age 18, and her self esteem, respectively. Thus we do not see, for instance, that pre-t_1 mothers systematically enter the labor force under better market conditions. (There is likewise no discontinuity if I plot parents' education, mother's age at first birth, or mother's labor force participation.) Thus, despite the evidence of the endogeneity of t_1 in Table 3, Figure 4 suggests that the use of t_1 to translate B_1 into K_1 is unlikely to be introducing a substantial additional source of bias into the comparison of these two populations of mothers.[35]

Evidence on Bias Captured in the Estimates of θ and γ

As initial evidence for the possible causality of first-birth timing on women's long-run wages, Figure 5a plots the experience profile of wages for post-t_1 mothers in the NLSY79, grouped by their relative timing of first birth, K_1.[36] Figure 5b plots the same for pre-t_1 mothers.

[35] In footnote 16, I raised the possibility that my estimates of δ are biased because women sort into a pre- or post-t_1 first birth based on the relative size of g_{0i} and g_{2i}, or that women with higher post-motherhood wage growth rates systematically choose a pre-t_1 first birth. Yet unless the variation in these parameters arises from factors that are completely uncorrelated with the observable characteristics discussed here, these results argue against either of these scenarios.

[36] This figure includes all mothers in the NLSY79 who are observed at least 20 years beyond t_1. If I plot

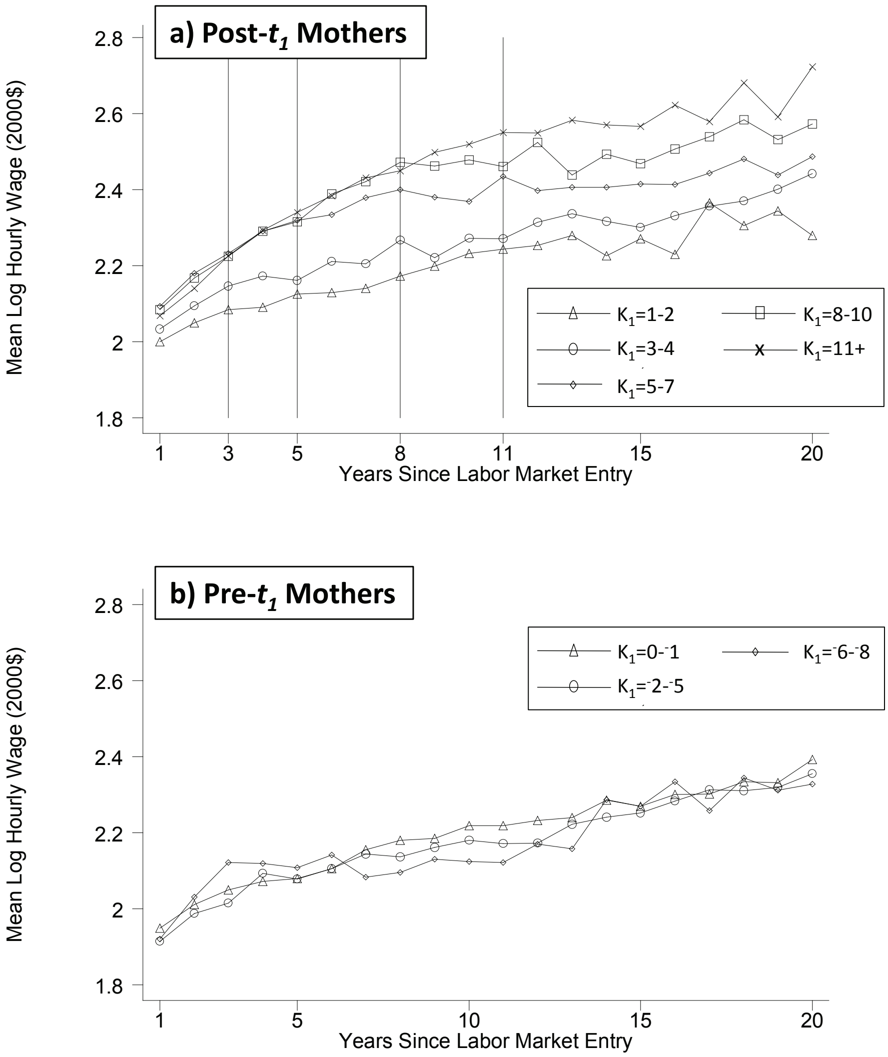

Figure 5: Wage Path by Timing of First Birth

The main pattern evident in Figure 5a is that among post-t_1 mothers, women's wages move roughly together before motherhood, and diverge sharply thereafter. For instance, consider women with $K_1 \geq 3$. Up till their third year in the labor market – before any have children – their wages move together (although women with $K_1 = \{3, 4\}$ are at a lower level). At that point, the first group reaches motherhood, and their growth rate stalls.

Figure 5a in more narrowly-defined education groups, in each I find the same clear pattern of a kink in the wage path after motherhood. Although each point in these figures captures data only for women with observed wages in the given year, changes in composition do not drive the pattern for post-t_1 mothers. For instance, if I plot average AFQT scores or education, including in each year only those women with an observed wage, there is no discontinuous drop at K_1. (As a measure of the composition of post-t_1 mothers present in the labor force over time, for all timing groups, average AFQT scores remain very stable.)

By comparison, the remaining groups continue to move together, diverging only after each following group begins to have children. (I see the same pattern if I plot average hours worked over time, or the proportion in the labor force; labor supply moves together before first birth, and diverges only as each group has their first child.[37])

Thus overall, we see that the raw wage path in Figure 5a looks strikingly similar to the wage pattern for Scenario #1 in Figure 2, where the difference in long-run wages between early and late mothers reflects a causal effect of the timing of first birth. In particular, we see that wages move together until each group has their first child, at which point the wage path for mothers kinks.[38] We do not see that wages diverge directly upon labor market entry, as anticipated by both Scenario #2 (endogenous timing) and Scenario #3 (unobserved heterogeneity), nor do we see that wages diverge *before* K_1, as in Scenario #4 (reverse causality).[39]

Furthermore, Figure 5a looks surprisingly similar to the piece-wise linear wage path depicted in Equation (1). First, we see that the rate of wage growth before first birth, g_1, is very similar across timing groups, suggesting that although g_1 may vary across women, it does not vary systematically with K_1. And second, tracing out an average wage path *after* first birth, g_2 looks surprisingly similar across timing groups as well. This in turn fits the description of Scenario #1, in which the return to delay is invariant because g_1 and g_2 are equal across the population, and variation in observed timing of first birth arises only from variation in taste for early motherhood.

[37]By contrast, and mirroring the pattern in Figure 5b, the labor supply patterns of all pre-t_1 mothers are very similar. I also find that for women with $K_1 > 0$, the disruption evident at K_1 – in both the wage path and labor supply – is peculiar to the *first* birth.

[38]By contrast, when lining up wages instead by age and grouping these two populations, Miller (2011) finds that wages diverge before first birth, although Wilde *et al.* (2010) do not.

[39]Since these wages reflect annual data, it is possible that we would not observe a kink that fell just before K_1. Furthermore, even in a reverse causality scenario, the kink in the wage path might line up K_1 if women can both perfectly anticipate when their wages will stall, and perfectly time their first birth accordingly. Note also that my OLS method described in Section 5.2 does not directly address this possible source of bias. Yet if I regress a woman's wage level at t_5 on her *subsequent* first-birth timing (among women with $K_1 > 5$), I find no "effect" of K_1 – thus no evidence of reverse causality – even though in each education group roughly 45 percent of these women have their first birth within the next two years.

In Figure 6 I next plot average AFQT scores by K_1, now showing all education levels.[40] As described in Scenarios #2 and #3, positive ability bias may drive the link between timing and wages if high-ability women systematically choose later first births. As evident in Figure 6, however, among post-t_1 mothers, for all but high school graduates, timing and AFQT scores are almost completely unrelated. Thus for women with a first birth after t_1, Figure 6 largely undermines the possibility that ability bias is driving the link between timing and wages.

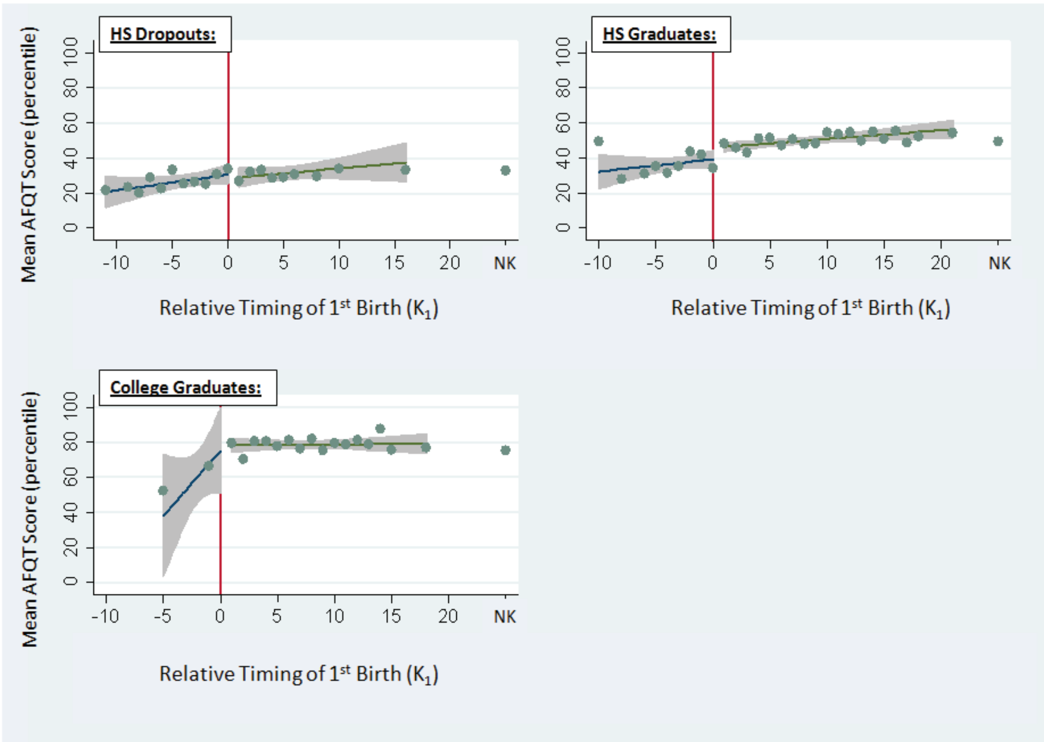

Figure 6: Mean AFQT Scores by Timing of First Birth

As the second test for bias, Table 4 reports the results of the series of regressions described in Section 5.2. Grouping women by their education level at t_1, in each panel I report the OLS-estimates of θ, γ, and δ (the coefficients on k_1, $k_1^{pt_1}$, and pt_1, respectively), as well as the coefficients on several of the key covariates added at the various stages of the identification process.

[40]This figure uses data for the NLSY79 cross-section sample who are observed at least 20 years beyond t_1 (including women with no kids, "NK"), including simple fitted lines and their 95-percent confidence intervals. The figures are very similar using my analysis sample.

Table 4: Estimates of the Effect of First-Birth Timing on Wages

Controls:	None (1)	Raw (2)	+ X_{a_i} + (3)	+ X_{ψ_i} (4)	+ X_{b_i} (5)	+ K_1^{pred} (6)	- w_0 (7)	By age (8)
High School Dropouts:								
k_1 (θ)	0.014	0.009	0.009	0.009	0.009	0.008	0.007	-
	(0.011)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	
$k_1^{pt_1}$ (γ)	0.026***	0.019	-0.009	-0.009	-0.011	-0.010	-0.011	-
	(0.008)	(0.012)	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)	
pt_1 (δ)	0.184**	0.232***	0.239***	0.223***	0.204***	0.355***	0.193***	-
	(0.090)	(0.081)	(0.080)	(0.079)	(0.078)	(0.100)	(0.078)	
Age at first birth ($\tilde{\theta}$)	-	-	-	-	-	-	-	-0.009
								(0.008)
Starting wage ($l(w_0)$)		0.085	0.100	0.081	0.034	0.043	-	0.021
		(0.066)	(0.066)	(0.065)	(0.071)	(0.074)		(0.075)
Total kids by t_{20}		-0.065***	-0.062***	-0.062***	-0.051**	-0.049**	-0.051**	-0.051**
		(0.021)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)	(0.023)
AFQT score (x10^1, in 1979)			0.018	0.014	-	-	-	-
			(0.012)	(0.012)				
Years off			-0.032**	-0.032*	-0.029	-0.029	-0.029	-0.010
			(0.016)	(0.017)	(0.018)	(0.018)	(0.018)	(0.014)
Total kids expected (at t_1)				0.011	0.008	0.001	0.009	0.006
				(0.023)	(0.023)	(0.022)	(0.023)	(0.023)
K_1^{pred} (at t_1)						0.024*	-	-0.003
						(0.013)		(0.012)
R^2	**0.02**	**0.31**	**0.35**	**0.36**	**0.44**	**0.44**	**0.43**	**0.42**
High School Graduates:								
k_1 (θ)	0.018***	0.017***	0.014***	0.014***	0.015***	0.015***	0.013***	-
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	
$k_1^{pt_1}$ (γ)	0.022**	0.013	-0.004	0.000	0.001	0.000	0.002	-
	(0.010)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	
pt_1 (δ)	0.113**	0.135***	0.135***	0.122***	0.116**	0.090	0.107**	-
	(0.051)	(0.050)	(0.047)	(0.047)	(0.047)	(0.058)	(0.048)	
Age at first birth ($\tilde{\theta}$)	-	-	-	-	-	-	-	0.009***
								(0.004)
Starting wage ($l(w_0)$)		0.294***	0.242***	0.246***	0.235***	0.235***	-	0.227***
		(0.045)	(0.046)	(0.045)	(0.045)	(0.045)		(0.045)
Total kids by t_{20}		-0.014	-0.015	-0.020	-0.022*	-0.022*	-0.024**	-0.026**
		(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
AFQT score (x10^1, in 1979)			0.043***	0.042***	-	-	-	-
			(0.006)	(0.006)				
Years off			-0.021**	-0.017	-0.015	-0.015	-0.015	-0.014
			(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Total kids expected (at t_1)				0.023*	0.020*	0.021*	0.017	0.023**
				(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
K_1^{pred} (at t_1)						-0.002	-	-0.007

Table 4 – Continued

Controls:	None	Raw	$+ X_{a_i}$	$+ X_{\psi_i}$	$+ X_{b_i}$	$+ K_1^{pred}$	$- w_0$	By age
						(0.007)		(0.006)
R^2	**0.02**	**0.16**	**0.24**	**0.24**	**0.29**	**0.29**	**0.27**	**0.28**
College Graduates:								
k_1 (θ)	0.035***	0.031***	0.032***	0.032***	0.035***	0.036***	0.035***	-
	(0.011)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)	(0.012)	
$k_1^{pt_1}$ (γ)	-0.028	-0.043	-0.049	-0.045	-0.009	-0.011	-0.017	-
	(0.061)	(0.059)	(0.059)	(0.059)	(0.056)	(0.057)	(0.064)	
pt_1 (δ)	-0.165	-0.118	-0.129	-0.121	0.037	0.032	-0.054	-
	(0.196)	(0.205)	(0.217)	(0.213)	(0.224)	(0.229)	(0.226)	
Age at first	-	-	-	-	-	-	-	0.031***
birth ($\tilde{\theta}$)								(0.012)
Starting wage		0.379***	0.348***	0.349***	0.366***	0.373***	-	0.353***
($l(w_0)$)		(0.094)	(0.099)	(0.099)	(0.111)	(0.111)	-	(0.110)
Total kids		-0.072	-0.046	-0.032	-0.037	-0.039	-0.018	-0.038
by t_{20}		(0.046)	(0.048)	(0.049)	(0.054)	(0.054)	(0.055)	(0.054)
AFQT score			0.014	0.015	-	-	-	-
(x10^1, in 1979)			(0.022)	(0.022)				
Years off			-0.015	-0.020	-0.026	-0.026	-0.015	-0.025
			(0.059)	(0.061)	(0.062)	(0.063)	(0.064)	(0.052)
Total kids				-0.028	-0.033	-0.035	-0.021	-0.032
expected (at t_1)				(0.040)	(0.041)	(0.041)	(0.041)	(0.042)
K_1^{pred} (at t_1)						-0.003	-	-0.001
						(0.018)		(0.017)
R^2	**0.06**	**0.28**	**0.30**	**0.31**	**0.37**	**0.37**	**0.33**	**0.36**

NOTES: This table lists the coefficients (and robust standard errors) from a weighted regression of Equation (4), controlling for an increasing number of individual characteristics (see footnote 29 for a listing of the controls included in all specifications but the first): "None" includes no controls; "Raw" estimates largely exclude individual characteristics; "$+X_{a_i}$" adds controls for AFQT and Rotter scores, years off, self-esteem, and expected education and work status at age 35; "$+X_{\psi_i}$" adds controls for total kids expected, gender-role attitudes, mother's age at first birth, and own age at first sexual intercourse; "$+X_{b_i}$" adds controls for family background (*e.g.*, race, religion, mother's education, and characteristics of the household at age 14); "$+K_1^{pred}$" controls for a woman's prediction of her own timing of first birth, made at approximately t_1. In Column (3) and Column (4) I replace missing values with their means because I do not yet want to capture any unobserved heterogeneity reflected in the incidence of missing data. When I control for family background, X_{b_i}, I now substitute for these means with missing value indicators. At this stage I also interact AFQT scores by race, hence its exclusion from the table from Column (5) forward. If I run the analysis without the survey weights, the pattern of how the estimates change with the inclusion of these controls is very similar. Column (7) reruns the Column (5) specification excluding starting wages. I use the Column (5) specification given my discussion on whether the inclusion of w_0 influences my interpretation of δ; if I instead use the Column (6) specification, the results are almost identical to reported those in Column (6). Column (8) reruns the Column (6) specification replacing k_1, $k_1^{pt_1}$, and pt_1 with age at first birth. Significance levels are indicated by * (at the 10% level), ** (at 5%), and *** (at 1%).

Column (1) of Table 4 reports these estimates before including any controls at all. (Those shown are from a specification excluding starting wage, but as with the comparison of Column (5) and (7), the inclusion of w_0 has little effect.) For all education levels we see a positive slope on K_1 for post-t_1 mothers ($\hat{\theta}$), as well as for pre-t_1 mothers ($\hat{\gamma}$) among both high school graduates and dropouts, the two education levels with appreciable numbers of the latter. Thus at first glance, there appears to be a similar link between K_1 and long-run wages for both populations of mothers, albeit with a trend break at $K_1 = 0$ ($\hat{\delta} > 0$).

Column (2) then reports my "raw" coefficient estimates – the partial correlations between K_1 and $l(w_{20})$ after controlling for factors that influence the distribution of w_{20} (*e.g.*, calendar year and location), as well as starting wages and some family characteristics. For college and high school graduates we still see a clearly positive link for post-t_1 mothers (very similar to the Column (1) estimates), but no clear link for high school dropouts. By contrast, we see that the link for pre-t_1 mothers is no longer significantly positive.

In Column (3) we see the influence on these estimates of controlling for X_{a_i} (*e.g.*, AFQT scores, highest grade expected, and years off), factors that may directly influence a woman's wage, and may likewise enter into her marginal benefit of delay. For $\hat{\theta}$, we see relatively little change. The coefficient for high school graduates drops by almost 20 percent, driven by the correlation between AFQT scores and K_1 evident in Figure 6, as well as by controlling for highest grade expected and years off. But for the other education levels, $\hat{\theta}$ is completely unchanged.

For γ, we instead see that controlling for X_{a_i} eliminates entirely the positive relationship between timing and wages observed among high school dropouts and graduates. This is driven in part by AFQT scores, but the more important element of X_{a_i} generating the positive correlation between timing and wages is years off.[41] By comparison, the estimates $\hat{\delta}$

[41] As seen in Column (3) of Table 3, among pre-t_1 mothers, those women with higher potential wages enter the labor market more promptly after B_1. Since the Column (3) results of Table 4 also show a link between years off and w_{20}, this suggests that years off acts as a proxy for some element of underlying labor market quality that is correlated with a woman's wages throughout her career. Yet note that among pre-t_1 mothers, years off is very highly (negatively) correlated with K_1 – perfectly so for the 68 and 49 percent of high school dropouts and graduates, respectively, who have their first birth by the end of the calendar year following their graduation year, and who do not return to school before entering the labor force. Thus when I control for years off, $\hat{\gamma}$ will be estimated off only the remaining women (the coefficient on years off is also being identified off of post-t_1 mothers). Yet this does not appear to drive my results. For instance, among the pre-t_1 high school graduates, the relationship between $l(w_{20})$ and years off (if I exclude $k_1^{pt_1}$) is very similar for the "perfectly correlated" subset as for the remaining women. But when I regress $l(w_{20})$ on $k_1^{pt_1}$

are very stable, thus the estimated level difference in w_{20} between women with a first birth just before versus just after t_1 are completely unchanged by controlling for X_{a_i} (as one would anticipate given the evidence in Figure 4).

In the next column I control for X_{ψ_i}, factors that may be correlated with taste for early motherhood. If, as in Scenario #3, the link between timing and wages arises from a correlation between ψ and factors that influence wage growth, controlling for X_{ψ_i} should drive the estimates $\hat{\theta}$ and $\hat{\gamma}$ towards zero. Instead we see that their inclusion has little effect.[42] When I next control for background characteristics, X_{b_i}, the R^2 rises throughout, but the estimates $\hat{\theta}$ and $\hat{\gamma}$ still remain unchanged. (By comparison, in each of these two steps the estimates $\hat{\delta}$ for high school graduates and dropouts fall slightly, decreasing in total by 13 to 14 percent of their initial "raw" estimates.)

Lastly I control for K_1^{pred}, a woman's prediction of K_1 based on her information set at t_1. (The term is undefined for pre-t_1 mothers, thus the Column (4) specification provides my "final" estimates for $\hat{\gamma}$ and $\hat{\delta}$.[43]) If, as in Scenario #2, variation in θ_i introduces positive bias because women facing a higher return to delay choose later first births, then controlling for K_1^{pred} should absorb this, driving the estimates $\hat{\theta}$ towards zero. Instead we see that the coefficients are very stable, even though K_1^{pred} is a good predictor of observed timing.[44] Furthermore, since K_1^{pred} is unrelated to w_{20} (even if I exclude all other individual controls), the *type* of women who expect to have kids later do not have higher wages.

Overall, in line with the evidence in Figure 5a, the results in Table 4 suggest that there is relatively little bias captured in the correlation between first-birth timing and long-run

excluding years off, the slope on $k_1^{pt_1}$ for the "perfectly correlated" subset is approximately the reverse of the previous coefficient on years off, whereas the coefficient on $k_1^{pt_1}$ for the remaining women is insignificantly different from zero. This suggests that the key relationship is between $l(w_{20})$ and years off.

[42] If I instead add X_{ψ_i} before X_{a_i}, the coefficients are again completely unaffected by their inclusion. Yet all of the elements of X_{ψ_i} are significantly correlated with timing for at least one education group, although among post-t_1 mothers, the link is more common with predicted than observed timing.

[43] Since K_1^{pred} is undefined for pre-t_1 women, when I include it in the regressions, to the extent that the intercept of the slope between K_1^{pred} and $l(w_{20})$ is non-zero, this will be loaded onto the coefficient for pt_1.

[44] Across education groups, K_1 increases by 0.3, 0.5, and 0.7 years for each 1-year increase in K_1^{pred}.

wages for women who remain childless at t_1 (θ).[45] Among these post-t_1 mothers, I only find evidence of bias for the high school graduates, where the 'raw' coefficient is somewhat inflated by positive ability bias.

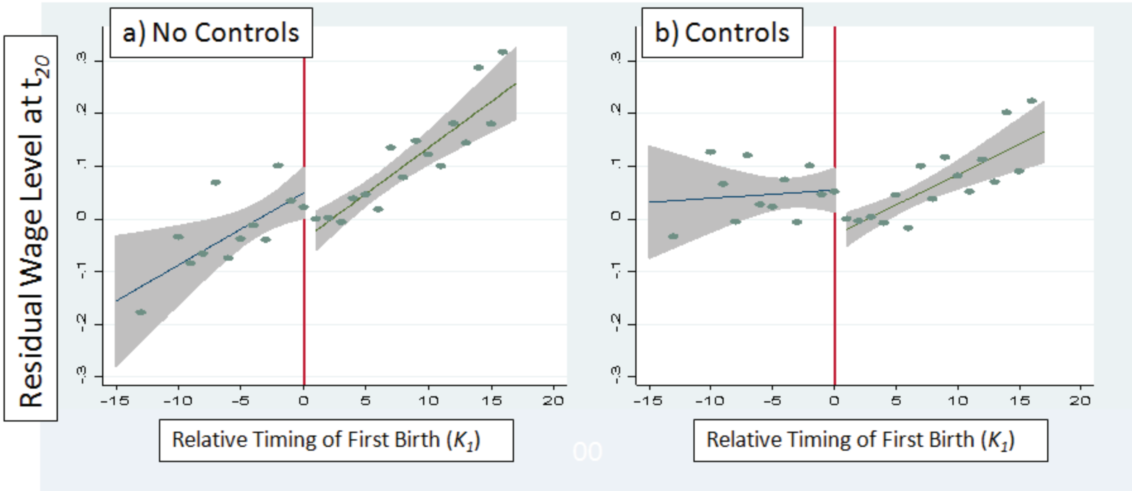

Figure 7: Mean Residual Wage Level at t_{20} by K_1

We can see this likewise in Figure 7, where I plot w_{20} by K_1 (for all education levels combined), first with no controls, and second residual of the full set of controls (exclusive of K_1^{pred} and years off).[46] In particular, we see that the positive slope for post-t_1 mothers ($K_1 > 0$) falls only slightly from Figure 7a to 7b. By contrast, and mirroring the results in Table 4, Figure 7 shows clearly that the initial positive link between K_1 and w_{20} for pre-t_1

[45] If I follow Miller (2011) in using miscarriages and contraceptive failures as instruments for first-birth timing, I also find no evidence that the OLS estimates of θ are biased upwards. In particular, as in Herr (2008), I use Miller's method in the following, amended way: (*i*) I do not use her third instrument (time to pregnancy) for the reasons discussed in Wilde *et al.* (2010), (*ii*) I interact the incidence of miscarriages and contraceptive failures with the age at which they occur, and (*iii*) I limit my focus to women who conceive their first birth within marriage, since, as shown in Herr (2008), the incidence of these instruments is only uncorrelated with individual characteristics among this population, thus these health outcomes are clearly not randomly distributed for the remaining women. Rerunning my fully-controlled specification on this population, the OLS and IV estimates of θ for high school graduates are 0.018 (s.e. 0.005) and 0.019 (s.e. 0.014), respectively (the latter only significant at the 20-percent level). For college graduates, the estimate rises from 0.030 (s.e. 0.009) to 0.044 (s.e. 0.025), significant at the 1- and 10-percent levels, although the difference is insignificant. (Because so few pre-t_1 mothers are married at conception, I cannot use this method to estimate γ and δ; the same holds for θ for high school dropouts.)

[46] Figure 7 includes simple fitted lines for pre- and post-t_1 mothers, and their 95-percent confidence intervals. The figure is identical in pattern if I use only high school graduates. The specification used in calculating the residual wages for Figure 7b excludes the three key regressors, k_1, $k_1^{pt_1}$, and pt_1. I also exclude K_1^{pred} because its inclusion complicates the interpretation of $\hat{\delta}$, and, building off the discussion in footnote 41, I exclude years off because of its strong correlation with K_1 for pre-t_1 mothers. (One might otherwise worry that the reason the slope on K_1 for these women disappears from Figure 7a to 7b is because I am controlling for something so strongly correlated with K_1.) If one instead includes years off, the slope for pre-t_1 mothers is even flatter, and slightly negative.

mothers is driven by differences in individual characteristics.

Consider this in light of the results in the existing literature, which consistently finds that the raw correlation between timing and wages is biased upwards. For instance, relative to OLS estimates, using fixed effects, Amuedo-Dorantes and Kimmel (2005) find a smaller estimated return to delaying first birth beyond age 30, and using IV, Miller (2011) finds a smaller return to delay among women in their 20s and 30s. Geronimus and Korenman (1992) and Hotz *et al.* (2005) likewise find a smaller negative impact of a teen birth using fixed effects and IV approaches, respectively. The evidence in Figure 7 suggests that the bias reported in the existing literature is driven by the spurious raw correlation between timing and wages for pre-t_1 mothers. Thus in comparison to my findings of relatively little bias in the link between timing and wages for post-t_1 mothers – the majority of women – this suggests that the positive bias reported in the existing literature is largely a byproduct of combining these two groups of mothers.

6.2 Gauging the Correct Measure of "Timing"

I now return to the question posed at the start of the paper, of the appropriate way to measure timing of first birth, given my focus on its effect on a woman's wage level. Consider, for instance, high school graduates, the education level with the greatest balance between pre- and post-t_1 mothers. Table 4 shows a clearly positive relationship between K_1 and long-run wages for post-t_1 mothers, $\hat{\theta}$, yet given the imprecision of the slope estimate of $\hat{\gamma}$, we cannot reject that the slopes are the same on both sides of t_1. Namely, the return to delay may be the same for both populations of mothers.

Yet returning to Figure 5b, the raw wage data for pre-t_1 mothers suggests that there is no relationship between K_1 and a woman's subsequent wage path. In stark contrast to the evidence in Figure 5a for post-t_1 mothers, we see that the wage path for each group of pre-t_1 mothers is very similar. (This same contrast in the pattern for pre- and post-t_1 mothers is evident in Figure 7b.) And note that this lines up with the results of the teen birth literature

– most of whom will be pre-t_1 mothers – that finds little to no effect on women's wages (*e.g.*, Geronimus and Korenman, 1992, and Hotz *et al.*, 2005).

Thus, given this evidence, I conclude that $\gamma = g_0^{'} = 0$. Namely, there is no linear link between first-birth timing and a woman's long-run wage level for those women who have their first birth *before* they enter the labor force. Consequently, I conclude that the appropriate measure for gauging the effect of first-birth timing on women's wages is the 'career timing' of first birth, the point in a woman's career in which children are first present.

Now consider the implications of this conclusion in light of the existing literature's grouping of these two populations of mothers. Ignoring for the moment the difference between K_1 and age at first birth, from the pattern in Figure 7b it is clear that the average slope across the full range of K_1 is going to underestimate the slope for post-t_1 mothers, θ. Accordingly, when I replace my measures of timing with age at first birth (see Column (8) of Table 4), I find a "return" of 0.9 and 2.0 percent, for high school and college graduates, respectively. Thus, consistent with my prediction in Section 2.2, when fertility timing is measured in terms of age, the estimated return to one year of delay understates θ by 40 and 15 percent, respectively.[47] Also as predicted, the difference is smallest for college graduates, because of the low proportion of pre-t_1 mothers.

6.3 The Returns to Delayed First Birth

<u>The Benefit of Fertility Delay if $K_1 > 0$</u>

Given the conclusions in Sections 6.1 and 6.2, consider now the economic implications of the results in Table 4. First, we see that among women who remain childless at labor market

[47]Note that the estimates of the return to delay in Chandler *et al.* (1994) and Miller (2011), 2.1 and 3.0 percent, respectively, are much higher than $\hat{\theta}$ for my pooled sample (1.6 percent). (Only these papers report the return to a single year of delay.) This is driven exclusively by their sample selection criteria, which capture a subset of "high type" mothers – only married and working full time for the former, and those with a first birth between 21 and 33 for the latter. Building a sample consistent with Chandler *et al.*, I find that the return to delay measured in terms of "career timing" is 60 percent larger than when measured in terms of age. By contrast, using a sample consistent with Miller, I find an identical return whether measuring in terms of age or career timing. Yet consistent with my discussion in Section 2.2, this is because only 2 percent of this sample has a pre-t_1 first birth.

entry, for all but high school dropouts, there is a clear positive return to additional fertility delay.[48] For high school graduates, one year of delay leads to 1.5 percent higher wages 20 years after labor market entry, corresponding to 3.2 percent of total wage growth over this 20-year stretch. By comparison, the return to a one-year delay of first birth for college graduates is 3.6 percent, or a much larger 6.9 percent of total wage growth by this point.

The Benefit of a pre-t_1 First Birth

Yet for high school dropouts and graduates, the positive estimates of δ suggest that in a range around t_1, it may be strictly better to have one's first birth *before* entering the labor force. We see this visually in Figure 7b in the level drop in the residual of w_{20} from those with a first birth just before t_1 to those with a first birth just after.

More precisely, given its definition, $\delta \equiv (g_0 - g_2)t_{20}$, $\hat{\delta} > 0$ suggests that $\bar{g_0} > \bar{g_2}$.[49] Namely, pre-t_1 mothers experience a higher average wage growth rate than the post-motherhood growth rate observed for post-t_1 mothers. Furthermore, one can define the "cross-over" point, τ^x, as the point beyond which a woman must delay her first birth, in order to experience greater wage growth by t_{20} than if she had her first child before t_1: $g_1\tau^x + g_2(t_{20} - \tau^x) = g_0 t_{20}$. Rearranging and substituting terms, $\tau^x = \delta/\theta$.

Based on my coefficient estimates, the cross-over point for high school graduates is strikingly far along: $\hat{\tau}^x = 7.9$ years. Since among the post-t_1 mothers, 56 percent have their first child before this, my results imply that possibly the majority of high school graduate women would have higher long-run wages if they had their first child before entering the labor force, rather than after. The same holds for high school dropouts, since for these women I find no clear evidence of a return to delay beyond t_1.

[48] Another possibility is that the effect of children on a woman's wage path is the same, regardless of the timing of her first, but that the cost evolves over time. In that case, when I observe the sample at t_{20}, I am capturing these women at different stages of this process. To consider this possibility, I compare the estimates of θ when measured at t_{15} and t_{20} for women with $K_1 \leq 12$ (approximately 85 percent of each education group). I find that the estimates $\hat{\theta}$ are very stable: $\hat{\theta}_{20}$ is 12 percent smaller than $\hat{\theta}_{15}$ for college graduates, and the two are identical for high school graduates.

[49] For high school graduates, using $\bar{t}_{20} = 21.5$, my estimates suggest that $g_0 - g_2 = 0.005$. Since this value is less than $\hat{\theta}_{HS}$, where $\theta = g_1 - g_2$, these results do not suggest that $g_0 > g_1$.

Is a Pre-t_1 First Birth Really Better?

The results above are admittedly surprising, especially given the magnitude of the estimated benefit of a pre-t_1 first birth. Can having one's first child before labor market entry really be (that much) better for women with less than a college degree?

Notice, first, that because my main specification controls for starting wage (w_0) – following my model in Section 2 – this means that my results primarily speak to the effect of fertility timing on wage *growth* rather than levels. Furthermore, we see from Table 2 that among both high school graduates and dropouts, women with their first birth just before t_1 enter the labor market at lower average wages than those with a first birth just after. Since my results may therefore only reflect that their wages *catch up* over time, they may not imply that pre-t_1 mothers in fact "do better."

Yet the lower starting wages of pre-t_1 mothers are fully explained by background characteristics, and thus excluding w_0 from my specification has little effect on $\hat{\delta}$ (see Column (7) of Table 4). Furthermore, if I replace w_{20} with either own lifetime earnings from t_1 to t_{20}, or total household earnings – thus capturing information on the husbands these women marry – mothers with their first birth just before t_1 experience significantly better outcomes than those with their first birth just after. (Because pre-t_1 mothers in fact work more during these 20 years, the return in terms of own earnings is larger than in terms of wages.[50]) The same holds if I instead measure earnings from age 20 to 40, since we may believe that women make their fertility and career choices in terms of age rather than career timing.[51]

The observation that pre-t_1 mothers work more from t_1 to t_{20} raises another possible

[50]For instance, for high school graduates I find that women with their first birth just before t_1 earn $83,000 more from t_1 to t_{20} than those with a first birth just after (in year-2000 dollars, significant at the 1-percent level), which, compared to average total earnings of $406,000, reflects a 20 percent return. (This total lines up with, for instance, the average earnings of $29,700 observed for those working at least 1500 hours in 2000, since this population works at least 1500 hours for an average 13.4 years.)

[51]The results reported here reflect rerunning the Column (4) specification on each of these dependent variables, separately for high school dropouts and graduates. The magnitude of the return to a pre-t_1 first birth, measured as a percentage of the mean, is smaller for household than own earnings. For instance among high school graduates, women with their first birth just before t_1 experience 5 to 6 percent higher household earnings, whether measuring from age 20 to 40, or from t_1 to t_{20}. The only instance in which $\hat{\delta}$ is not significantly positive is when regressing household earnings from 20 to 40 for high school dropouts.

complication. Note that the key distinction between pre- and post-t_1 mothers is that the former have at least one of their children *before* they enter the labor force.[52] If having a child early induces women to work more, as found in Hotz *et al.* (2005), this may be part of the mechanism of the return to a pre-t_1 first birth. But if these women simply take their "post-first-birth labor force gap" *before* t_1 – allowing them to then work more between t_1 and t_{20} – I may simply be mis-measuring the relevant window of potential experience.

To consider this, I rerun the main specification using the wage level approximately 20 years after a woman completes school. Thus I now focus on the same length of time for all women, where pre-t_1 mothers have chosen to front-load some of their parenting time before working. In contrast to the period from t_1 to t_{20}, I now find that pre-t_1 women work less during these 20 years.[53] I also find that the estimate $\hat{\delta}$ for high school dropouts is cut in half, suggesting that mis-measurement of the relevant window of potential experience may explain much of the estimated return at this education level. But for high school graduates, $\hat{\delta}$ instead rises slightly. Thus for these women, I again find evidence suggesting the benefit of a first birth just before labor market entry rather than just after.[54]

It is uncertain, however, whether an earlier first birth provides better outcomes on other dimensions, especially for the child – particularly if this would require a pre-marital birth.[55] I leave to future work this question of the mechanism of the benefit of a pre-t_1 first

[52] The magnitude of δ is not driven by mothers who complete their childbearing before t_1 (30 and 40 percent of pre-t_1 high school dropouts and graduates, respectively); controlling for this lowers $\hat{\delta}$ by less than 5 percent for both education levels.

[53] Measuring from t_1 to t_{20}, pre-t_1 mothers work on average 14 and 7 percent more hours, for high school dropouts and graduates, respectively. By comparison, in this 20-year stretch both groups of pre-t_1 mothers work on average 11 to 12 percent less. (For pre-t_1 mothers who return to school after B_1 (before t_1), I measure these 20 years from the year before B_1.)

[54] One might also wonder if differences in sample selectivity may influence my estimates. Yet if I recreate Table 1 separately for only high school dropouts or graduates, the included and excluded women are very similar among both the "just before" and "just after" groups.

[55] By my simple utility model in Section 3, because these women have both higher earnings *and* enjoy earlier parenthood, these pre-t_1 mothers should have strictly higher utility than early post-t_1 mothers. A more realistic model incorporating factors surrounding marriage, however, may suggest that these pre-t_1 mothers are worse off in other dimensions, such as the probability of raising their child within marriage or the quality of their spouse. As we see in Appendix Table A-1, women with a first birth just before t_1 are much less likely to be married at birth than those with a first birth just after. One might therefore hypothesize that pre-t_1 mothers spend fewer years married, or are married to lower-earning spouses, and that this drives their higher labor supply from t_1 to t_{20}. If so, their higher long-run wages may reflect a

birth, as well as the mechanism for the positive return to delay among women who instead have their first child after entering the labor force.

Does θ Vary?

As a last point, I consider whether θ varies over time or across women. In Section 2.1 I suggested that the return to fertility delay may arise from the postponement of the interruption of motherhood further into one's career, when the underlying wage path has started to plateau. This suggests that the return should fall with time, as the gradient of the wage profile becomes less steep. Furthermore, since the wage path is steeper among the more educated (Murphy and Welch 1992), the return to delay may increase with schooling, or with any other factor that leads to a steeper wage profile.[56]

I first test for a declining return to delay by including the quadratic of k_1 in Equation (4). Consistent with the linear pattern in Figure 7b, I find no evidence of a decreasing return among either high school or college graduates. When we consider that the vast majority of women have their first birth during their first decade in the labor force, a range in which the standard wage profile remains steep, this result may seem less surprising.

Testing for whether θ varies with education or ability, I find mild evidence of variation across the full population, but no evidence of additional variation within education levels. In particular, consistent with my results in Table 4, in the pooled sample I find that θ is

"marriage" effect, rather than a "timing" effect. Yet I find no evidence to support this. For instance, despite a discontinuous jump at $K_1 = 0$ in the percent married at first birth, there are no corresponding jumps in either average spouse quality, or average spouse earnings from K_1 to t_{20}, both of which are controlled for throughout my specifications.

[56] In footnote 1 I also raised the possibility that the return to delay varies with the *reason* for that delay. To the extent that I can test this, I find only limited evidence of variation in θ. In particular, I can (imperfectly) isolate three of the possible reasons listed there: (a) "I wasn't married yet," (c) "it took me longer to get pregnant than I expected," and (f) "I got pregnant unexpectedly." For (a), I compare whether the return to delay is systematically different for women who conceive immediately after marriage, to approximate the population with pent-up demand for a (marital) first birth. Likewise, for (c), I use as the comparison women who were actively trying to get pregnant at least one year before they conceived, and for (f), I use the sample of women who report that they did not want to get pregnant when they conceived their first child. For both (c) and (f), I find no evidence that the returns are different (and where sample sizes allow sufficient precision, the estimated returns are very similar). Only for (a) do I find any evidence that the return varies. Among high school graduates, the return is significantly smaller (and may be zero) for women who may be delaying only because they remain unmarried.

increasing by 0.4 percentage points for each year of schooling at t_1 (significant at the 5-percent level). By comparison, I do not find that the return is significantly increasing in AFQT scores. Furthermore, within high school and college graduates, I also find no evidence of additional variation in the return, by either education or ability.

7 Conclusion

The first conclusion of this analysis is that the existing literature has mis-estimated the return to fertility delay by its focus on a woman's age at first birth. When one considers the mechanisms by which fertility timing may influence a woman's wage path, this suggests that the critical factor is her experience level at first birth, not her age. Furthermore, this leads to the insight that the effect may vary for women with a first birth before versus after they enter the labor force. In the analysis above, I show that the relevant measure is the career timing of a woman's first birth, the point in her career when children are first present.

One result that emerges from this conclusion is that there is relatively little bias captured in the raw correlation between the timing of a woman's first birth and her long-run wage level for women who enter the labor market before motherhood. This result is evident in Figure 5a, where women's wages move together before motherhood, and stall directly thereafter. I also find that by grouping these women with pre-t_1 mothers, the existing literature has been underestimating the return to delay for this population. Furthermore, I find that the return is surprisingly invariant across women, especially within education levels, which helps explain why we see no evidence of endogenous sorting in response to this large return to delay.

The more striking results that emerge from this re-definition of timing are the updated estimates of the return to delay for pre-t_1 mothers. Although I find no linear link between timing and wages among this population, I find a clear benefit of having one's first birth before labor market entry rather than after. This suggests that the existing literature's almost universally-reported positive estimates have concealed the benefit of an *earlier* first birth for this population of mothers. Thus it may be that the effect of children on women's wages largely arises from the career disruption of the shift into motherhood, rather than the presence of children themselves.

References

[1] Amuedo-Dorantes, Catalina and Jean Kimmel. 2005. The Motherhood Wage Gap for Women in the United States: The Importance of College and Fertility Delay. *Review of Economics of the Household*, 3: 17-48.

[2] Becker, Gary S. 1985. Human Capital, Effort, and the Sexual Division of Labor. *Journal of Labor Economics*, 13(1, part 2): S33-S58.

[3] Blackburn, McKinley L., David E. Bloom, and David Neumark. 1993. Fertility Timing, Wages, and Human Capital. *Journal of Population Economics*, 6(1):1-30.

[4] Bloom, David E. 1986. Fertility Timing, Labor Supply Disruptions, and the Wage Profiles of American Women. In *Proceedings of the Social Statistics Section*, pp. 49-63.

[5] Buckles, Kasey. 2008. Understanding the Returns to Delayed Childbearing for Working Women. *American Economic Review, Papers and Proceedings*, 98(2): 403-07.

[6] Cameron, Stepehn V., and James J. Heckman. 1993. The Non-Equivalence of High School Equivalents. *Journal of Labor Economics*, 11(1): 1-47.

[7] Caucutt, Elizabeth M., Nezih Guner, and John Kowles. 2002. "Why Do Women Wait? Matching, Wage Inequality, and the Incentives for Fertility Delay." *Review of Economic Dynamics*, 5(4):815-855.

[8] Chandler, Timothy D., Yoshinori Kamo, and James D. Werbel. 1994. Do Delays in Marriage and Childbirth Affect Earnings? *Social Science Quarterly*, 75(4): 838-853.

[9] Cigno, Allesandro, and John Ermisch. 1989. A Microeconomic Analysis of the Timing of Births. *European Economic Review*, 33: 737-360.

[10] Geronimus, Arline T., and Sanders Korenman. 1992. The Socioeconomic Consequences of Teen Childbearing Reconsidered. *Quarterly Journal of Economics*, 107(4): 1187-1214.

[11] Hamermesh, Daniel S., and Jeff E. Biddle. 1994. Beauty and the Labor Market. *The American Economic Review*, 84(5): 1174-1194.

[12] Happel, S.K., J.K. Hill, and S.A. Low. 1984. An Economic Analysis of the Timing of Childbirth. *Population Studies*, 38(2): 299-311.

[13] Heckman, James J. and James R. Walker. 1990. The Relationship Between Wages and Income and the Timing and Spacing of Births: Evidence from Swedish Longitudinal Data. *Econometrica*, 58(6): 1411-1441.

[14] Herr, Jane Leber. 2008. Does it Pay to Delay? Decomposing the Effect of First Birth Timing on Women's Wage Growth. Unpublished manuscript.

[15] Hotz, Joseph V., Susan Williams McElroy, and Seth G. Sanders. 2005. Teenage Childbearing and Its Life Cycle Consequences: Exploiting a Natural Experiment. *The Journal of Human Resources*, 40(3): 683-715.

[16] Light, Audrey, and Manuelita Ureta. 1995. Early-Career Work Experience and Gender Wage Differentials. *Journal of Labor Economics*, 13(1): 121-154.

[17] Miller, Amalia R. 2011. The Effects of Motherhood Timing on Career Path. *Journal of Population Economics*, 24(3): 1071-1100.

[18] Moffitt, Rober. 1984.Mincer, Jacob and Haim Ofek. 1982. Interrupted Profiles of Fertility, Labour Supply and the Wages of Married Women: A Complete Life-Cycle Approach. *The Review of Economic Studies*, 51(2): 263-278.

[19] Mullin, Charles H., and Ping Wang. 2002. The Timing of Childbearing Among Heterogenous Women in Dynamic General Equilibrium. NBER Working Paper #9231.

[20] Murphy, Kevin M., and Finis Welch. 1990. Empirical Age-Earnings Profiles. *Journal of Labor Economics*, 8(2): 202-229.

[21] Murphy, Kevin M., and Finis Welch. 1992. The Structure of Wages. *The Quarterly Journal of Economics*, 107(1): 285-326.

[22] Perry, Cynthia D. 2005. How do Female Earnings Affect Fertility Decisions? Working Paper.

[23] Rindfuss, Ronald R., Larry Bumpass, and Craig St. John. 1980. Education and Fertility: Implications for the Roles Women Occupy. *American Sociological Review*, 45(3): 431-447.

[24] Schmidt, Lucie, "Murphy Brown Revisited: Human Capital, Search, and Nonmarital Childbearing among Educated Women," Working Paper, 2007.

[25] Stange, Kevin. 2011. A Longitudinal Analysis of the Relationship Between Fertility Timing and Schooling. *Demography*, 48(3): 931-56.

[26] Taniguchi, Hiromi. 1999. The Timing of Childbearing and Women's Wages. *Journal of Marriage and the Family*, 61(4): 1008-1019.

[27] Topel, Robert H. and Michael P. Ward. 1992. Job Mobility and the Careers of Young Men. *The Quarterly Journal of Economics*, 107(2): 439-479.

[28] Waldfogel, Jane. 1998. Understanding the "Family Gap" in Pay for Women with Children. *Journal of Economic Perspectives*, 12(1): 137-156.

[29] Walker, James R. 1995. The Effect of Public Policies on Recent Swedish Fertility Behavior. *Journal of Population Economics*, 8(3): 223-251.

[30] Wilde, Elizabeth Ty, Lily Batchelder, and David T. Ellwood. 2010. The Mommy Track Divides: The Impact of Childbearing on Wages of Women of Differing Human Capital. NBER Working Paper #16582.

Appendix

NOT FOR PUBLICATION

A Calculating the Marginal Benefit of Delay

Building on the utility maximization model introduced in Section 3.1, to solve for a woman's marginal benefit of fertility delay, we must first solve for the production function of earnings, $y = f(K_1)$. If I first make the simplifying assumption that labor supply is constant throughout the lifecycle ($h_t = 1$ for all t), the net present value of lifetime earnings is

$$f(K_1) = \int_T w_t h_t e^{-rt} dt = \int_0^{K_1} w_0 e^{g_1 t} e^{-rt} dt + \int_{K_1}^{T} w_0 e^{g_1 K_1} e^{g_2(t-K_1)} e^{-rt} dt,$$

where r reflects the discount rate and T the length of a woman's career. (As written, this assumes that T is constant across all women.) This, in turn, solves to

$$f(K_1) = \frac{w_0(g_2 - g_1)}{(g_1 - r)(g_2 - r)} e^{(g_1-r)K_1} + \frac{w_0}{(g_2 - r)} e^{(g_2-r)T} e^{(g_1-g_2)K_1} - \frac{w_0}{(g_1 - r)}. \tag{A-1}$$

The marginal benefit of delay will then equal $f'(K_1)/f(K_1)$. Taking the linear approximation of Equation (A-1), $MB(K_{1i}) = \theta_i - m_{1i}K_{1i}$, where $m_{1i} \geq 0$. (The slope on K_{1i}, $m_{1i} \equiv \theta_i \frac{g_{2i}-r}{e^{(g_{2i}-r)T}-1}$, is non-negative, regardless of the relative size of g_{2i} and r.) Note that although θ appears in both the numerator and denominator of K_1^*, for reasonable values of g_2, r, and T, the term in the numerator dominates.

If I instead assume that h_t varies by the presence and age of children, but not by the timing of the first:

$$h_t = 1 \qquad \text{for } t < K_1 \text{ and } \; t > K_1 + \Lambda, \qquad \text{and}$$
$$h_t = (1 - \lambda) \qquad \text{for } t \in \{K_1, K_1 + \Lambda\},$$

the production function of earnings, $f(K_1)$, is now equal to

$$f(K_1) = \int_0^{K_1} w_0 e^{g_1 t} e^{-rt} dt + \int_{K_1}^{K_1+\Lambda} (1-\lambda) w_0 e^{g_1 K_1} e^{g_2(t-K_1)} e^{-rt} dt + \int_{K_1+\Lambda}^{T} w_0 e^{g_1 K_1} e^{g_2(t-K_1)} e^{-rt} dt,$$

$$= \Big[\frac{w_0}{(g_1-r)} - \frac{w_0}{(g_2-r)}\Big(\lambda(e^{(g_2-r)\Lambda}-1)+1\Big)\Big] e^{(g_1-r)K_1} + \frac{w_0}{(g_2-r)} e^{(g_2-r)T} e^{(g_1-g_2)K_1} - \frac{w_0}{(g_1-r)}.$$

As expected, this expression reduces to Equation (A-1) if $\lambda = \Lambda = 0$. Without calculating the full linearization of the corresponding marginal benefit function, I can show that the intercept term is again a function of θ_i:

$$= \theta_i \frac{\Big(e^{(g_{2i}-r)T}-1\Big) - \lambda(g_{1i}-r)\Big(e^{(g_{2i}-r)\Lambda}-1\Big)}{\Big(e^{(g_{2i}-r)T}-1\Big) - \lambda\Big(e^{(g_{2i}-r)\Lambda}-1\Big)}.$$

B Sample

This analysis uses the sample of women from the 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). The survey began in 1979 with 6,283 14- to 22-year old women, who by 2008 had reached the ages of 44 to 51. The largest groups lost across time are the military and supplemental white/poor samples, which were dropped from the NLSY79 in 1984 and 1991, respectively. Because the earliest year used for "t_{20}" in my analysis is 1994, this means that these two samples are largely excluded. (Throughout this appendix I refer to career years by number, for instance, t_{17} for the 17th career year; for the purposes of this appendix, t_{20} will therefore reflect the 20th career year, and "t_{20}" will reflect the six-year window used for defining w_{20}.) Exclusive of these two groups, we observe 89 percent of the remaining sample through at least age 40, or approximately the end of the childbearing years. Of these, 83 percent had at least one child by their last year observed, when the median woman was 47.

My sample builds from the NLSY79 cross-sectional sample, plus the supplemental minority samples. Because of the inclusion of these oversamples, throughout the analysis I apply sampling weights that were built specifically for my final sample by the Center for Human Resource Research at the Ohio State University.[1]

[1] When I calculate statistics for a population broader than my sample, I use the 1979 sampling weights.

C Variable Definitions

To define t_1, the year of labor market entry, I first establish a woman's "graduation year," the year in which she completed continuous schooling. (Under certain circumstances I allow gaps in a woman's education, as long as upon her return she completed more years of schooling at a sufficient pace.[2]) Starting from June of that year, I then search for the first twelve-month period in which she worked at least 1000 hours.[3] I define that year — the second calendar year in that 12-month stretch — as t_1, her first year in the labor market.

I define a woman's 20th-year wage, w_{20}, using any wage for years t_{19} through t_{24}, centered on t_{21} or t_{22}.[4] Given this approach, I observe w_{20} for 77 percent of the NLSY79 women observed through age 40. I define a woman's starting wage as her wage in her first or second year in the labor force; of women who ever work, I observe starting wages for 90 percent of all women observed through age 40.[5]

In my final sample, by number, women from the supplemental sample make up 41 percent of high school dropouts, 38 percent of high school graduates, and 21 percent of college graduates. Applying the appropriate weights, these women instead represent 17, 16, and 7 percent of each education group, respectively.

[2]I allow a 1-year gap if, after returning to school, her reported education level increased. (Many respondents report school attendance without their completed education level increasing thereafter.) I also allow up to a 2-year gap if upon return she stayed in school for at least 2 years, completed at least 2 years of schooling in that timeframe, and completed it at a rate of at least 0.65 school years per calendar year (*e.g.*, finishing 2 school years in 3 years but not 4). Lastly, I allow up to a 4-year gap if during that time she never worked more than 1000 hours per year, and when she returned to school she completed at least 1 year of full-time, or 2 years of part-time, schooling. For women who drop out of high school, but later complete a GED, I define the graduation year based on the timing of the former.

[3]When determining the last year of school versus the first year of work, in some instances of part-time (post-high school) education, I count her as already in the labor force if she was simultaneously working at a rate of 1000 hours or more per year.

[4]For all wage data in a given year t I use the wage for the current job, or if not currently working, for the job that corresponds with a time closest to the middle of the calendar year (based on the date the job ended). In some cases of implausible values I use an alternate wage (corresponding to different jobs worked), or correct values that appear to be typos. In some cases where neither gives me a reasonable alternative, I treat it as a missing value. I also exclude all hourly wages below \$2 and above \$200, both in year 2000 dollars, with the upper bound lower for earlier survey years. (Wages are converted to year-2000 dollars using the Consumer Price Index for all urban consumers.) Because for each woman the 20th year falls beyond 1993, when the survey became biennial, I use only data reported for the survey years in building w_{20}. For those who entered the labor market in an odd year, I use the first available wage for their 22nd, 24th, and 20th career year (in that order), and for those who entered in an even year, I use their 21st, 23rd, or 19th.

[5]Of those lacking a starting wage, 71 percent entered the labor force before 1978; the initial survey did not capture retrospective wage data.

For other variable definitions:

- The NLSY79 collected information on fertility expectations – when a woman expected to have her next child, and how many (more) children she expected – in 1979, 1982 to 1986, and biennially from that point forward. For women with $K_1 > 0$, I build K_1^{pred} using data for t_1 or the year before. For all women, I measure total children expected at approximately t_1.

- To address assortative mating, I build a measure of a woman's spouse "quality" – as a potential proxy for her own "quality" – using the average of her husband's residual annual earnings (for all years married from 1978 through 2007), after controlling for the following factors: her age, race, education, AFQT score, any pre-marital children, SMSA and region of residence, and calendar year. I specifically want to capture her husband's earnings residual of *her* characteristics, rather than *his*, because, for instance, the observation that a high school-educated woman is married to a college-educated man may reflect her "high quality." I also control for children that predate the given couple because I do not want to capture the marriage market penalty for existing kids (*e.g.*, women remarrying after a divorce).

- I calculate average spouse resources for the years from K_1 till "t_{20}" (starting at 1978, the first year with available earnings data), using husband's earnings (when married or reported for a partner), plus child support and alimony payments as reported, but otherwise including 0s for years in which a woman is unmarried.

- The gender-attitudes measure is the sum of the responses to a set of questions concerning family- and gender-role attitudes. For instance, a woman was asked in 1979 how strongly she agreed with the following statement: "A wife who carries out her full family responsibilities doesn't have time for outside employment," or "A working wife feels more useful than one who doesn't hold a job." I align the seven questions so that a higher score reflects a more conservative view, and sum these responses.

- For father's occupation and women's expected occupation at age 35, I group detailed occupations into six broad categories: traditionally male professional jobs, traditionally female professional jobs, sales, clerical, services, and blue-collar jobs. Because of sample-size issues, for father's occupation I also combine various of these categories, depending on the woman's education level.

- For women with a high school diploma or less, I separately distinguish Hispanics from African-Americans; for college graduates I group minorities.

D Sample Restrictions

$K_1 > 17$:

Given that my dependent variable is a woman's wage level at "t_{20}," I limit the sample to women with $K_1 \leq 17$, since among the latest mothers, I do not want to capture a sample that is influenced by the rate at which women return to work after K_1, which may be correlated with wages. Among the cross-section sample who I observe for at least 10 years after first birth, 86 percent return to work within 36 months, and 91 percent within five years. With the bulk of my sample using wages from t_{21} or t_{22}, by limiting myself to women with $K_1 \leq 17$, most women will be back to work by this point.[6]

Age at First Birth > 35:

Among women from the NLSY79 cohort, evidence suggests that as women approached the end of their 30s and the oncoming fertility decline, some chose to have a first child under lesser circumstances. To show this, Figure A-1 plots, by age at first birth, the proportion of women who were married at the conception of their first child, married by the birth of that child, and among the latter who were married, their husband's earnings that year.[7]

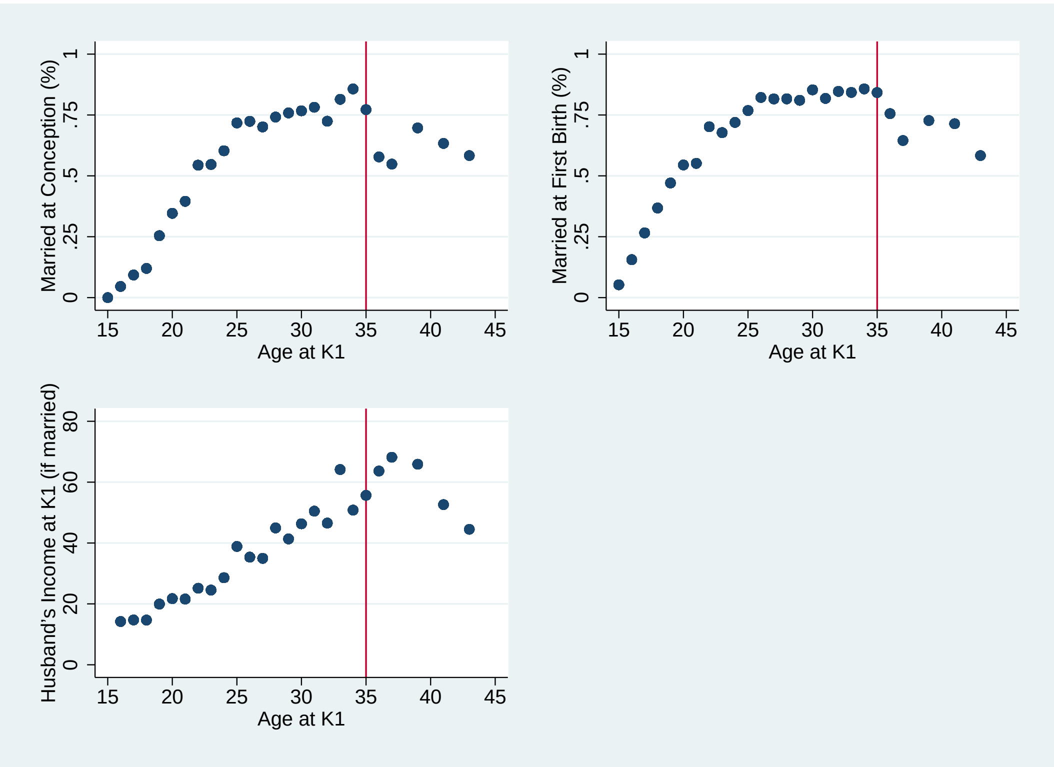

Figure A-1: Marriage Outcomes by Age at First Birth

In the first panel, after age 35 we see a discontinuous drop in the proportion of women married at the conception of their first child. Comparing women with their first birth between

[6]In addition, in building w_{20} I use only those wages that reflect a point greater then 36 months after a woman's first child was born.

[7]Age is measured in months, rounded to the closest integer. Figure A-1 includes all mothers observed through age 40.

the ages of 30 and 35, to those with a first birth after 35, 78 versus 61 percent are married (significantly different at the 1-percent level).

Although the proportion of women who are married by the *birth* of that child is higher, the same level drop appears after age 35. Furthermore, these differences are apparent within each education level. For instance, although a larger 11 percent of college graduates have their first birth after age 35 (compared to 5 percent overall), college graduates over 35 are 11 percentage points less likely to be married at birth (significant at the 1-percent level).

For women married at first birth, the last panel of Figure A-1 plots husband's average earnings that year. As expected, we see that husband's earnings rise with her age (and thus, on average, with his). In the late 30s, however, we see that the positive slope reverses. Thus, even among women who were married, it appears that later mothers may have 'settled' for lower-earning spouses in order to have a child. All of these results are consistent with the predictions of the "spousal search" model in Schmidt (2007).

The evidence in Figure A-1 suggests that the older mothers from the NLSY79 cohort may possess some characteristic that was penalized in the marriage market. If that same characteristic influenced their labor market outcomes – consider the literature on the relationship between earnings and beauty (*e.g.*, Hamermesh and Biddle 1994) – later mothers may likewise have had lower wages because of this underlying characteristic.

The data suggest this is true, but only clearly so among the college graduates (the largest group of older mothers). To show this, Figure A-2 plots college graduates' mean residual starting wage (w_0), wage growth to t_3, wage level at approximately t_5 (limited to women with their first birth at least one year later), and wage level at "t_{20}" (after all had had a first child).[8]

As seen in the first panel of Figure A-2, women who went on to have a first birth after

[8]Using all college graduates observed to age 40 with a first child by t_{17}, these values reflect the residual of a regression of the given dependent variable on race, education at t_1, AFQT scores interacted by race, highest grade expected at age 18, and location at t_1/t_5/"t_{20}."

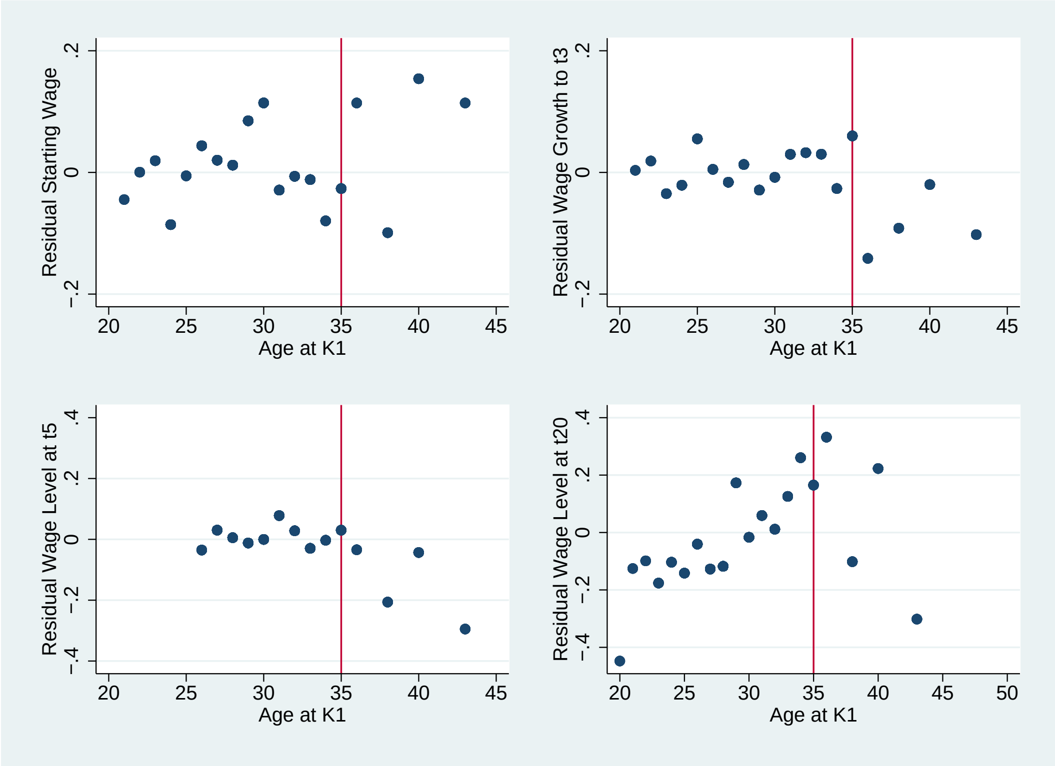

Figure A-2: Wage Outcomes by Age at First Birth

age 35 showed no systematic difference in starting wages. Yet the next two panels show that their wage growth lagged behind over time. In comparison to women with a first birth in their early 30s, they had lower wage growth by t_3, and likewise lower wages at t_5 (significant at the 5- and 1-percent levels, respectively).

The last panel of Figure A-2 plots the residual wage level at "t_{20}," after all of these college graduates had reached motherhood. Consistent with my results, we see an initial clear rise by age at first birth (reflecting the return to fertility delay), but this pattern reverses in the late 30s. The late mothers, who exhibited lower wage growth *before* motherhood, likewise have clearly lower wages afterwards.

In combination, this evidence suggests that these women have some underlying characteristic which is driving them to both have a late first birth *and* experience lower wage growth. This leads to the observation that these late mothers have lower long-run wages, but not because the late fertility timing is *driving* these lower wages. Based on this evidence, I have therefore excluded from my analysis those NLSY79 women with a first birth after age 35. If I were to instead include these women, my regression results in Section 6 would suggest a much lower return to delay for college graduates, driven by the spuriously low wages of these late mothers.

Given these two restrictions, my final sample of 2,567 mothers reflects 74 percent of all mothers observed till at least age 40.[9] Appendix Table A-1 reports an expanded set of summary statistics for this sample.

E Influence of Sample Selection Criteria

<u>Women Not Observed to "t_{20}":</u>

The low proportion of women captured in my sample among the earliest mothers is worrisome if these women are systematically different from later pre-t_1 mothers (see Table 1). Table A-2, which explores the reasons why excluded women are lost from the sample, shows that 22 percent of these excluded women never enter the labor force. A much larger 71 percent are lost because we do not observe them through at least t_{19} or t_{20}, the earliest years used for "t_{20}." (By comparison, among women with $K_1 > 0$, at most 11 percent are lost because they are not observed sufficiently long beyond t_1.)

In practice, the necessity of observing women through at least t_{19}/t_{20} means that I lose all women with their first birth more than 15 years before t_1. This restriction is mild for women with at least a high school diploma at t_1, but for high school dropouts, among those who ever work, 7 percent have their first child before this point. And as discussed in Section 4.2, Table 1 also shows that among the earliest mothers, there are large differences in AFQT scores between those included and excluded from the sample (although these differences only hold among high school dropouts). Thus overall we see that my sample

[9] Some women are also dropped because of missing labor force or school attendance data. Since the NLSY79 only captured retrospective labor force participation back to 1975, I also drop the handful of women who completed their schooling, or had a first child, before 1974. Of mothers observed till age 40, the sample captures 63 percent of high school dropouts, 78 percent of high school graduates, and 69 percent of college graduates, with education measured at t_1 or when last observed (for those who never work). (I do not limit my sample to women observed till age 40, although only 1.4 percent do not meet this criterion.)

Table A-1: Other Summary Statistics

	All	Fertility Timing (K_1)						
		$\leq^- 7$	$^-7-^-4$	$^-3-0$	$1-3$	$4-6$	$7-9$	10+
Less than High School:								
Total kids expected at t_1	2.3	2.9	2.3	2.3	2.2	2.2	2.5	2.0
Predicted timing at t_1 (K_1^{pred})	5.4	-	-	-	4.8	5.4	5.9	6.2
Mother's age at first birth	20.2	19.4	19.8	20.1	19.8	21.0	20.9	20.4
Family attitudes at t_1	14.5	15.0	13.7	14.3	14.4	14.9	14.5	14.1
Age at first intercourse	15.9	15.5	15.9	15.6	16.0	16.7	15.7	16.9
Self esteem (in 1980)	-0.39	-0.59	-0.35	-0.42	-0.45	-0.21	-0.18	-0.48
% Married at K_1	61.6	56.2	45.0	58.5	63.8	72.7	57.4	85.0
% Yrs married (K_1 to "t_{20}")	66.2	56.3	58.4	69.7	59.1	71.8	72.1	84.8
Average spouse resources, K_1-"t_{20}" ('000, 2000$)	21.7	16.4	17.2	21.7	19.3	27.5	23.8	29.5
Total kids at "t_{20}"	2.5	3.2	2.6	2.7	2.5	2.0	2.0	1.9
High School Graduate:								
Total kids expected at t_1	2.4	2.8	2.3	2.3	2.4	2.4	2.5	2.4
Predicted timing at t_1 (K_1^{pred})	5.4	-	-	-	4.4	5.3	6.0	6.2
Mother's age at first birth	21.0	20.3	20.5	20.0	20.7	21.1	21.4	22.2
Family attitudes at t_1	13.9	14.0	14.0	14.0	13.9	13.9	14.3	13.4
Age at first intercourse	17.4	16.2	16.4	16.4	17.2	18.0	17.9	17.9
Self esteem (in 1980)	0.02	-0.12	-0.04	-0.05	0.02	0.01	0.13	0.04
% Married at K_1	76.6	53.3	55.8	51.7	77.9	85.5	88.5	81.8
% Yrs married (K_1 to "t_{20}")	75.6	64.1	62.7	59.6	73.0	81.3	83.8	83.1
Average spouse resources, K_1-"t_{20}" ('000, 2000$)	31.7	22.1	21.9	21.9	27.9	34.9	36.5	40.7
Total kids at "t_{20}"	2.2	3.1	2.5	2.5	2.5	2.2	2.0	1.8
College Graduate:								
Total kids expected at t_1	2.4	2.4	3.1	2.6	2.6	2.4	2.4	2.3
Predicted timing at t_1 (K_1^{pred})	4.8	-	-	-	3.2	4.7	5.3	5.8
Mother's age at first birth	23.6	18.9	21.9	21.2	24.4	22.9	24.3	23.9
Family attitudes at t_1	12.4	14.1	14.5	13.1	12.6	12.3	12.4	11.8
Age at first intercourse	18.7	14.1	17.7	18.9	19.0	19.2	18.5	18.0
Self esteem (in 1980)	0.42	0.40	-0.05	0.37	0.42	0.49	0.35	0.42
% Married at K_1	95.3	30.4	0.0	89.0	96.1	97.8	96.1	97.1
% Yrs married (K_1 to "t_{20}")	90.9	27.1	52.3	88.0	92.0	92.5	92.3	91.4
Average spouse resources, K_1-"t_{20}" ('000, 2000$)	55.8	7.3	34.4	40.1	61.4	51.7	61.3	56.6
Total kids at "t_{20}"	2.2	2.4	2.7	2.6	2.5	2.3	2.1	1.8

See notes to Table 2.

Table A-2: Reason Missing from Analysis Sample

Timing of First Birth (K_1):	$\leq^- 8$	$^-7 - ^-4$	$^-3 - 0$	$1 - 3$	$4 - 7$	$8 - 11$	≥ 12
Among excluded, reason: (%weighted)							
Never entered labor force:	21.8	0.0	0.0	0.0	0.0	0.0	0.0
$K_1 > t_{17}$ or age at $K_1 > 35$:	0.0	0.0	0.0	0.0	2.3	2.0	74.0
Missing w_{20} because:							
Not observed thru t_{19}/t_{20}	70.9	34.9	13.8	11.1	8.7	3.6	0.0
Did not work during "t_{20}"	6.9	52.0	47.3	48.4	59.0	65.3	21.1
Missing wages during "t_{20}"	0.3	13.1	38.9	40.6	30.0	29.1	4.9
Sample size:	**387**	**80**	**87**	**81**	**122**	**92**	**189**

NOTES: As in Table 1, this table focuses on the NLSY79 women excluded from my sample who I observe till at least age 40, and who have a child by their last year observed. The table reports percentages calculated using the original 1979 sample weights.

systematically loses women who take a long time to enter the labor force after having their first child, and that those women are in turn negatively selected.

This raises the question of the influence of this selection on my estimates of γ and δ. To consider this, I compare the starting- and 10th-year wages of pre-t_1 mothers, between those who I do and do not observe through t_{19}/t_{20}.[10] This lets me consider whether the women that I lose start at a systematically different wage level, or show a different wage pattern over their first 10 years in the labor force, as insight into whether their 20th-year wages are also likely to be systematically different.

For both high school dropouts and graduates, I find no statistical difference in starting wages.[11] (This holds whether or not I control for characteristics observable at t_1, thus the significant differences in AFQT scores do not translate into different starting wages.) For the subset of these pre-t_1 women for whom I observe a "t_{10}" wage, within each group I also find no difference in wage growth over this first 10 years. In combination, since starting and "t_{10}" wages of pre-t_1 mothers who do not reach t_{19}/t_{20} are statistically equivalent to those who do, this suggests that their "t_{20}" wages would likewise be similar. Thus it is unlikely that my estimates of γ or δ are affected by the systematic exclusion of women who I do not

[10] I build "t_{10}" wages using any observed wage for years t_7 through t_{12}, starting as close to t_{10} as possible. Among the pre-t_1 women who ever work, I observe a "t_{10}" wage for 87 percent of high school dropouts and 90 percent of high school graduates.

[11] I do not consider college graduates, where the number of pre-t_1 women is small.

observe sufficiently long to reach "t_{20}."

Women with No Observed Wage During "t_{20}":

Consider now the low proportion of women captured in the sample among those with $K_1 \geq 12$ (see Table 1). In Table A-2 we see that this is driven primarily by women with a first birth beyond t_{17} or after age 35. In Table 1 we also see that the excluded women have systematically higher education and starting wages, and these differences persist within high school graduates (but not college graduates). This difference arises from the fact that late mothers with some college education at t_1 are more likely to have their first child after 35.

The greater concern among later mothers are the women who are missing from the sample because they never worked during the 6-year stretch used for "t_{20}." In most timing groups, roughly 50 to 65 percent of those excluded are lost for this reason. Among women with a first birth between $K_1 = \{-3, 11\}$, the remaining 30 to 40 percent excluded are lost because I do not observe a wage within this 6-year stretch, even though these women worked at some point during that timeframe.

This again raises the question of whether this exclusion influences my coefficient estimates. To consider this, for women with $K_1 > 0$, I compare the women excluded from my sample to those included on the following four dimensions: (*i*) AFQT scores, (*ii*) starting wages, (*iii*) pre-motherhood wage growth to approximately t_5, and (*iv*) wage level after first birth. (For women with $K_1 \leq 0$, I run the same comparison for all but the third factor.)

The first three tests let me consider whether the women lost from my sample are systematically positively or negatively selected, in terms of ability, starting wage (w_0), and pre-birth wage growth (g_1). (This last test lets me consider whether the incidence of missing data is correlated with K_1, since selection will only influence my estimate of θ if this incidence is in turn correlated with K_1.) The fourth test lets me consider whether heterogeneity in θ influences the probability that women remain working after children; one would suspect that the women who face an especially large drop in their wage growth after K_1 are more likely to leave the labor force.

Starting with the set of women with their first birth by t_{17} and age 35, whom I observe through at least t_{23}/t_{24} – the end of the range used for "t_{20}" – for women with $K_1 > 0$ I find no evidence of either positive or negative selection out of my sample in any education level. I find no statistical differences in AFQT scores, whether I consider only the set of women who never work during "t_{20}," or the full set of women with no observed w_{20}. I likewise find no difference in starting wages and no difference in pre-K_1 wage growth to "t_5."[12]

When I use any post-K_1 wage to re-estimate θ, however, for both high school dropouts and graduates I find that $\hat{\theta}$ for the excluded women is significantly smaller than the estimate for the included women.[13] (This suggests that those women who are missing from my sample may experience a *smaller* drop in wage growth at K_1.) Yet for both education levels, the coefficient estimate $\hat{\theta}$ for all women combined is statistically indistinguishable from the estimates in Table 4. By comparison, for college graduates the estimate $\hat{\theta}$ for the missing women is instead significantly *larger* than the estimate for the included women, but again the combined coefficient is statistically equivalent to the Table 4 estimate. In combination, these results suggest that my estimates $\hat{\theta}$ are largely unaffected by the exclusion of women for whom I lack w_{20}.

Turning last to the pre-t_1 women, among both high school dropouts and graduates, I find no differences between those for whom I do and do not observe a "t_{20}" wage, either in terms of ability, starting wage, or the alternate long-run wage level. Thus, just as with the

[12] I build a "t_5" wage using any pre-K_1 wage for t_3 through t_7 (as close to t_5 as possible); the sample therefore captures the vast majority of women with $K_1 \geq 4$. (Of all post-t_1 mothers, 74 percent of high school graduates and 82 percent of college graduates have $K_1 \geq 4$; I do not consider high school dropouts because only 33 percent have their first birth this late.)

[13] To test for differences in long-run wage growth among pre-t_1 mothers, and in post-K_1 wages among post-t_1 women, I build an alternate post-K_1 wage using data for any year between t_9 and t_{34} (exclusive of the "t_{20}" period). Using this alternate wage lets me consider whether the women who are unobserved during "t_{20}" have a systematically different θ, although note that I cannot include in this test those women who never return to work. Among post-t_1 mothers with missing w_{20}, I observe a post-K_1 wage for 86, 74, and 60 percent, respectively, across the education levels. The proportion is smallest for college graduates in largest part because they enter motherhood later, and so fewer have returned to work by their last survey year. Among pre-t_1 mothers with missing w_{20}, I observe this alternate long-run wage for roughly 90 percent of both high school dropouts and graduates.

case of women who do not reach t_{19}/t_{20}, I conclude that my estimates of γ and δ are unlikely to be affected by the exclusion of these women with missing w_{20}.

F The Possible Endogeneity of Education at t_1

Since throughout my analysis I group women by their education level at t_1, this raises a potential endogeneity issue if education and first-birth timing are jointly determined. For instance, if some women get unexpectedly pregnant while in school, and that pregnancy derails their education, then a woman's education level at t_1 may be endogenous to fertility timing for those pre-t_1 mothers.[14]

In practice, I find that surprisingly few pre-t_1 mothers were in school when they conceived their first child, and among those who were, the majority completed more schooling after the birth, suggesting that such a pregnancy need not permanently suspend a woman's schooling. In particular, among NLYS79 pre-t_1 mothers, only 13 percent of those who enter the labor market as high school dropouts, and 29 percent of mothers who start work with more education, conceived while in school. Furthermore, among that subset, 63 percent went on to get more education after the birth, 48 percent completing more schooling before t_1. (Among all who were part way through either high school or college when they gave birth, 41 and 21 percent, respectively, completed the given degree before entering the labor force.[15]) In combination, consistent with research finding at most weak negative effects of a teen birth on education (Geronimus and Korenman [1992] and Hotz *et al.* [2005]), this evidence suggests that the education level of pre-t_1 mothers is surprisingly unrelated to the timing of their first birth.

Yet education may also be endogenous to *intended* first-birth timing: women who want

[14]This suggests that some pre-t_1 women may be grouped in a lower education level than their 'potential,' thus if this type of endogeneity holds, when I compare women with a first birth just before and after t_1, this may drive upwards the average outcome for the pre-t_1 mothers in those education groups. Yet if those women who have – and *maintain* – an unplanned pregnancy while in school are *negatively* selected, this will introduce bias in the opposite direction.

[15]The former percentage excludes those who complete only a GED by t_1. (As a comparison, of all NLSY79 women who complete at least some college education, 48 percent complete the degree.)

an early birth may choose less schooling (Blackburn *et al.*, 1993). To test for such a link, I compare a woman's highest grade wanted (as reported in high school) to her mother's age at first birth (a strong correlate of own age at motherhood[16]), grouping women into "potential education" levels based on their mother's education.[17] Among the daughters of the most educated mothers ("college types," based on the average completed education of this group), I find no link between mother's age at first birth and own educational expectations. Thus for this group, I find no evidence suggesting that those who want an earlier first birth choose less schooling.

Yet in the other two education levels, I find a significantly positive relationship. Thus among "high school" or "some college" *types* – the average education in these two groups is 12.7 and 14.0 years, respectively – this evidence suggests that those with a higher taste for early motherhood may systematically choose less schooling. Although throughout this analysis I group women with 12 to 15 years' schooling, this endogeneity may mean that some early mothers are grouped into an education level below their 'potential,' and may therefore have higher wage growth than the average woman in their observed education group. To the extent that these women choose "baby first," this may lead me to overestimate the benefit of a pre-t_1 first birth. If they instead work first, and thus choose an early post-t_1 first birth, their higher wage growth, combined with their early first birth, will dampen my estimate of θ. Yet, as seen in Figure 4, we see no evidence of an over-representation of "high potential" women among either pre-t_1, or early post-t_1, mothers.

[16]The correlation between own and mother's age at first birth is 0.13, 0.22, and 0.23 among the daughters of mothers with less than, exactly, and more than a high school diploma, respectively. (Although the NLSY79 collects expected year of first birth, its correlation with expected schooling could reflect the reverse direction of causation.)

[17]I group women accordingly to consider this comparison in more homogenous subsamples, since one may worry that own educational expectations and mother's age at first birth are correlated because both are driven by mother's education. All of the following results also hold, however, if I group women by their father's education, and throughout these regressions, I control for both parents' education.